Analysis and protease-catalysed synthesis of sucrose alkanoate regioisomers

Lie, Aleksander

Publication date:
2014

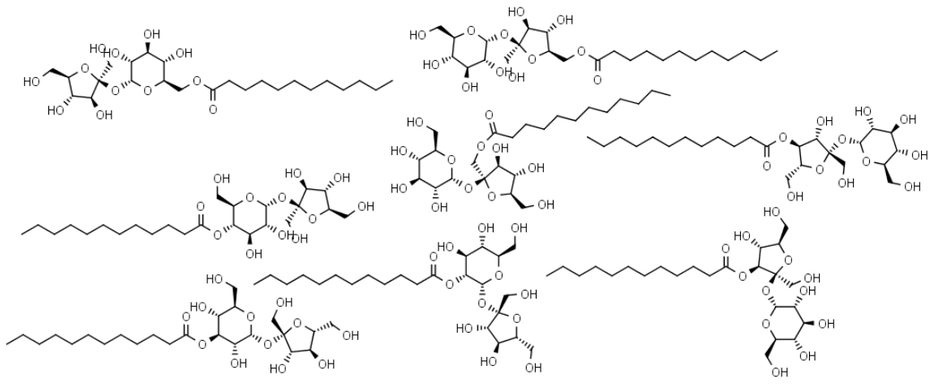
Document Version
Accepted author manuscript, peer reviewed version

ANALYSIS AND PROTEASE-CATALYSED SYNTHESIS OF SUCROSE ALKANOATE REGIOISOMERS

ALEKSANDER LIE

ISBN: 978-87-93100-54-1

ANALYSIS AND PROTEASE-CATALYSED SYNTHESIS OF SUCROSE ALKANOATE REGIOISOMERS



ALEKSANDER LIE

AALBORG UNIVERSITY
DENMARK

DEPARTMENT OF BIOTECHNOLOGY, CHEMISTRY AND ENVIRONMENTAL ENGINEERING
AALBORG UNIVERSITY - SOHNGÅRDSHOLMSVEJ 57 - 9000 AALBORG - DENMARK

Analysis and protease-catalysed synthesis of sucrose alkanoate regioisomers

Aleksander Lie

PhD Thesis

Printed May 2014

Section of Biotechnology

**Department of Biotechnology,
Chemistry and Environmental Engineering**

Aalborg University

CONTENTS

PREFACE

The work presented herein has been conducted in association with the Bioprocess Technology group at the Section of Biotechnology, Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University (AAU), Aalborg, Denmark.

The scientific journey of this PhD has been challenging, instructional, frustrating, eventful, delightful and rewarding. I have had the opportunity to attend interesting conferences and visit wonderful locations, and living and working in Aalborg and Denmark has been a great experience.

My gratitude and appreciation go first and foremost to my supervisor, Associate Professor Lars Haastrup Pedersen, who has been a great mentor and excellent help with research and academic writing. Our collaboration has been rewarding and productive.

A heartfelt thank you to the other members of our research group during my time as a PhD Fellow: Associate Professor Niels T. Eriksen, for constructive criticism and discussions during group meetings; Lab Technicians Tanja Bergmann and Charlotte Sten for invaluable assistance with practical matters; and fellow PhD Camilla Kær Mørkholt for scientific and moral support.

I am grateful to Associate Professor Reinhard Wimmer for the assistance in performing and analysing the results of NMR spectroscopy, and to Associate Professor Allan Stensballe for performing mass spectrometry analyses.

Also, thank you to the students that have been associated with our small team of researchers during this period, for simple questions, challenging questions, teaching and learning opportunities, and for providing variation to the working days.

And finally, thank you to the rest of the staff at the Section of Biotechnology, for any and all help and assistance, and for providing a warm and positive working environment.

Aalborg, January 2014

Aleksander Lie

SUMMARY

The aims of the presented research were to develop quantifiable methods for reversed-phase high-performance liquid chromatography analysis of sucrose alkanoate regioisomers and to investigate the activity and regioisomeric distribution in the biocatalytic esterification of sucrose with vinyl laurate in DMF using serine proteases and a metalloprotease.

A broad range of elution strategies for the chromatographic analysis of sucrose alkanoate regioisomers was systematically investigated using design of experiments strategies and statistical and multivariate analysis and modelling. Efficiency evaluation of the elution strategies, in terms of the resolution metric *general resolution deviation for multiple peaks* (RD_m) and analysis time, showed that the best results were offered by step-down gradient elution for sucrose caprate and isocratic elution with increased flow for sucrose laurate. Step-down gradient elution of sucrose caprate offered improvements in separation by 13-38 % and reductions in analysis time by 3-9 % compared to isocratic elution. For sucrose laurate, isocratic elution with increased flow showed improvements in separation of up to 29 % and reductions of 6-34 % in analysis time compared to isocratic elution. The optimised elution strategies for the sucrose alkanoates resulted in the separation of regioisomers with resolutions adequate for quantification, and the regioisomers were successfully identified and assigned by nuclear magnetic resonance spectroscopy.

The acylation of sucrose with vinyl laurate acyl donor in the hydrophilic, aprotic solvent *N*,*N*-dimethylformamide with or without protease biocatalyst resulted in the formation of sucrose laurate, with 2-*O*-lauroyl sucrose the most abundant monoester regioisomer synthesised. The serine protease ALP-901 produced the highest yield after 48 hours reaction time (53 %), while the reaction with thermolysin achieved the overall highest yield (63 %) after 6 hours, with only monoesters synthesised. The highest concentration of 2-*O*-lauroyl sucrose observed was 23.7 mM after 24 hours in the thermolysin-catalysed reaction. The detected appearance of the sucrose laurate regioisomers largely corresponded to the apparent rates of formation, based on monoester concentration and regioisomeric distribution after 48 hours, and 2-*O*-lauroyl sucrose was among the first regioisomers to appear in all investigated reactions. The concentration of sucrose laurate was observed to decline with reaction times above 6 hours in the reaction with thermolysin due to oligoester formation, and declining concentration with increasing reaction time was also observed for some specific sucrose laurate regioisomers in the reactions with other enzyme formulations. Differences in regioisomeric distribution after 48 hours appeared partly to stem from differences in overall reaction rates, as the different reaction conditions resulted in similar distributions after different reaction times. The esterification of sucrose with vinyl laurate with no protein in the reaction mixture appeared to be catalysed by the presence of aluminosilicate molecular sieves in the reaction medium. Mass spectrometry analysis of sucrose laurate product confirmed the molecular mass.

SAMMENDRAG

Målene for den præsenterede forskning var udvikling af kvantificerbare metoder for omvendt-fase højydelses væskekromatografisk analyse av sukrosealkanoatregioisomerer og undersøgelse av aktivitet og regioisomerfordeling i biokatalytisk esterdannelse fra sukrose og vinyllaurat i DMF ved brug af serinproteaser og en metalloprotease.

Et bredt udvalg av elueringsstrategier for den kromatografiske analyse af sukrosealkanoatregioisomerer blev systematisk undersøgt ved hjælp af eksperimentel-design-strategier og statistisk og multivariat analyse og modellering. Effektivitetsvurdering af elueringsstrategierne, baseret på målet for opløsningsevne *generelt afvig i opløsningsevne for flere toppe* (RD_m), og analysetid, viste at de bedste resultater blev opnået med nedtrinsgradienteluering for sukrosekaprat og isokratisk eluering med øget volumetrisk hastighed for sukroselaurat. Nedtrinsgradienteluering for sukrosekaprat gav forbedringer i separation på 13-38 % og reduktioner i analysetid på 3-9 %, sammenlignet med isokratisk eluering. For sukroselaurat gav isokratisk eluering med øget volumetrisk hastighed forbedringer af separation med op til 29 % og reduktioner på 6-34 % i analysetid, sammenlignet med isokratisk eluering. De optimerede elueringsstrategier for sukrosealkanoater resulterede i separering af regioisomerer med opløsningsevne tilstrækkelig for kvantificering, og regioisomerene blev identificeret og anvist med held ved hjælp af kernemagnetisk resonansspektroskopi.

Acylering af sukrose, med vinyllaurat som acyldonor, i den hydrofile, aprotiske solvent *N*,*N*-dimethylformamid med og uden proteasebiokatalysator, resulterede i dannelse af sukroselaurat, med 2-*O*-lauroylsukrose som den monoester syntetiseret i højest forekomst. Serinproteasen ALP-901 gav højest udbytte efter 48 timer reaktionstid (53 %), mens reaktionen med sinkproteasen thermolysin opnåede det generelt højeste udbytte (63 %) efter 6 timer, med kun monoestere syntetiseret. Den højeste koncentration af 2-*O*-lauroylsukrose observeret var 23,7 mM, efter 24 timer i den thermolysin-katalyserede reaktion. Detekteret forekomst av sukroselauratregioisomerene korresponderede i stor grad med de tilsyneladende dannelsesrater, baseret på monoesterkoncentration of regioisomerfordeling efter 48 timer, og 2-*O*-lauroylsukrose var blandt de første regioisomerer der forekom, i alle reaktioner. Forskelle i regioisomerfordeling efter 48 timer fremstod som delvis forårsaget av forskelle i de overordnede reaktionshastigheder, eftersom de forskellige reaktionsbetingelser resulterede i lignende fordelinger efter uens reaktionstider. Esterdannelse fra sukrose og vinyllaurat, uden protein i reaktionsblandingen, fremstod som katalyseret af tilstedeværelsen af aluminiumsilikatmolekylærsier i reaktionsmediet. Den molekylære masse af sukroselauratprodukt blev bekræftet ved massespektrometrisk analyse.

SYMBOLS & ABBREVIATIONS

$\in$	element of
$\propto$	proportional to
α	significance level
ε_r	relative static permittivity
μ	dipole moment
σ	half width of ideal peak at inflection point (standard deviation)
τ	Kendall rank correlation coefficient
$\bar{\omega}$	slope of peak width
A	peak area
A_s	asymmetry factor
a_w	water activity
ACN	acetonitrile
ANOVA	analysis of variance
AOF	aggregate objective function
CAD	charged aerosol detection
CCC	circumscribed central composite
CCD	central composite design
CMC	critical micelle concentration
COSY	correlated spectroscopy
DOE	design of experiments
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
e	mathematical constant (Euler's number), $e = \sum_{n=0}^{\infty} \frac{1}{n!} \approx 2.71828$
ELSD	evaporative ligh-scattering detection
FCC	face-centred cube
Fru	fructose
Glc	glucose
h	peak height
HCT	high-capacity ion trap
HPLC	high-performance liquid chromatography
HSQC	heteronuclear single quantum coherence
logP	partition coefficient
LC-MS	liquid chromatography - mass spectrometry
M	injected mass of analyte
MALDI	matrix-assisted laser desorption/ionisation
MeOH	methanol
MLR	multiple linear regression
MS	mass spectrometry
NMR	nuclear magnetic resonance
OFAT	one-factor-at-a-time
p	probability
pH	acidity/basicity of an aqueous solution, pH = $-\log_{10}[H_3O^+]$
pK_a	logarithmic acid-ionisation constant
PLS	partial least-squares
r	sample correlation coefficient

R^2	coefficient of determination
$R^2_{adj.}$	adjusted coefficient of determination
R_s	resolution
RD	resolution deviation
RD_m	general resolution deviation for multiple peaks
RID	refractive index detection
RMSD	root-mean-square deviation
RMSE	root-mean-square error
RP-	reversed-phase (chromatography)
RSD	relative standard deviation
t_0	dead time
T_B	boiling point
t_R	retention time
t_R'	adjusted retention time
TFA	trifluoroacetic acid
UV	ultraviolet (light)
w_b	peak width at baseline

Part I

CONTENTS OF PART I

1 INTRODUCTION

Carbohydrates are a considerable reserve of renewable starting materials for industries utilising carbon-containing molecules. In the past decades they have attracted continual interest for the synthesis of a variety of compounds, such as surfactants or polymerisable derivatives (Queneau et al., 2004). Carbohydrate chemistry is complex due to the polyfunctionality of these polyhydroxy aldehydes or ketones, many products can result from simple reactions and they lead to a host of derivatives, such as sugar alcohols, sugar acids, deoxysugar and amino sugars (Robyt, 1998). Of these ubiquitous compounds, sucrose is the most abundant pure organic molecule produced at the industrial scale (above 125 Mt/a) from renewable resources such as sugar cane and sugar beet (Queneau et al., 2004).

Sugar fatty acid esters are non-ionic surfactants commonly applied in the cosmetic, food, detergent and pharmaceutical industries (Ferrer et al., 2005; Plou et al., 2002), as they are physiologically, dermatologically and biologically acceptable (Allen and Tao, 1999). The physical and chemical properties of sugar fatty acid esters depend on the saccharide group, fatty acid chain length, and both position and degree of esterification (Puterka et al., 2003), and these amphiphilic compounds can be synthesised with a wide range of different emulsifying, dispersing and sensory properties (Garti et al., 1999; Vulfson, 1993). Some sugar fatty acid esters have been shown to possess antibiotic, antitumoral and insecticidal properties (Ferrer et al., 2005; Okabe et al., 1999; Piao et al., 2006; Puterka et al., 2003).

Biocatalysis in organic solvents is today a well-established field of research and application for the synthesis of organic compounds. Enzyme-catalysed chemical transformations offer several advantages compared to conventional chemical catalysis and organic synthesis, as enzymes show great versatility in substrate acceptance, exhibit high activity and selectivity, function under mild conditions and are environmentally acceptable (Faber, 2004; Koeller and Wong, 2001). Ester synthesis can be catalysed by hydrolytic enzymes, and particularly lipases and proteases have been applied to the esterification of disaccharides with success in various pure and mixed organic media (Polat and Linhardt, 2001; Shi et al., 2011).

The scientific work for this thesis has been performed as part of a research project titled *Enzyme-catalysed production of carbohydrate-based biosurfactants and investigation of their antimicrobial properties*, funded by the Danish Council for Independent Research. The aims of the research were to develop quantifiable methods for reversed-phase high-performance liquid chromatography analysis of sucrose alkanoate regioisomers and to investigate the selectivity and reactivity of enzyme-catalysed syntheses of sucrose laurate. The effects of essential chromatographic conditions were systematically studied using design of experiments strategies. The influence of the conditions on the retention and separation of regioisomers in complex samples of sucrose alkanoates was determined by statistical and multivariate analysis and modelling of the experimental data, and separation and analysis time were optimised based on the results. The biocatalytic esterification of sucrose with vinyl laurate in *N*,*N*-dimethylformamide was investigated using a range of protease formulations. The reactions were sampled over time and the sucrose laurate regioisomer distribution analysed to assess and compare the activity and selectivity of the enzymes.

1.1 Sucrose alkanoates

Fatty acid esters of sucrose are available with varying emulsifying, dispersing and solubilising properties, which combined with low toxicity, low irritancy and high biodegradability, make them attractive for use in foods, cosmetics and pharmaceuticals (Baker et al., 2000; Gupta et al., 1983). They are stable at pH 5-8 and at temperatures up to 180 °C (Okumura et al., 2011; Ushikusa et al., 1990), and approval as a food additive (E473) in most parts of the world leads to the production of low-substituted sucrose alkanoates as food and cosmetics emulsifiers on an industrial scale (Queneau et al., 2004).

Sucrose shares the polyhydroxy substitution of all carbohydrates, meaning that many products can arise from simple reactions, resulting from different degrees of substitution (mono-, di-, oligo-, etc.) and from the different positions of the substituents in the sucrose molecule (Queneau et al., 2004). This makes regioselective esterification of sucrose challenging, whether using chemical or enzymatic catalysis. Chemical synthesis is usually performed at high temperatures, resulting in poor selectivity and side reactions giving coloured derivatives (Polat and Linhardt, 2001). Notably higher selectivity has been shown for enzyme-catalysed processes with milder reaction conditions (Shi et al., 2011).

In terms of physicochemical properties, the sucrose alkanoates show a general decrease in water solubility with increasing alkanoyl chain length and degree of esterification. Antimicrobial properties have been shown for the more water soluble monoalkanoates with alkanoyl chain lengths of 10-16 carbon atoms (as described in section 1.1.2).

1.1.1 Regioisomers

Compounds with identical chemical composition, but exhibiting positional differences in chemical structure, are known as regioisomers. The eight hydroxy groups of sucrose can all participate in ester formation, leading to eight possible regioisomers of sucrose monoalkanoates which are labelled according to the carbon atom numbering shown in fig. 1-1. Intramolecular hydrogen bonds and differences in conformational stability lead to variations in the properties of the sucrose hydroxy substituents, with 2-OH being the most acidic (Queneau et al., 2008), and the position of attachment of aliphatic chains to a disaccharide affects the self-organising properties of the molecule and the bonding and structural features of the material (Queneau et al., 2001).

The regioisomers of sucrose alkanoates have been shown to have different physicochemical properties. The critical micelle concentrations (CMC) of regioisomers of specific sucrose esters have been reported to vary significantly, as the CMCs of 1′-*O*-myristoyl sucrose and

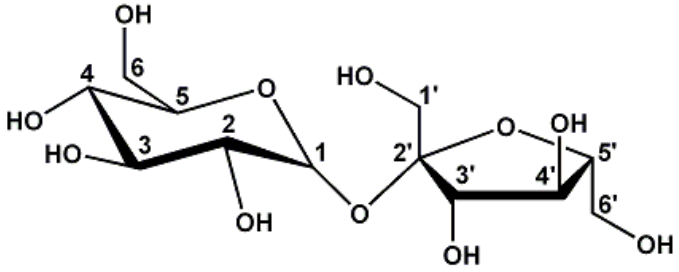


Figure 1-1. Sucrose structure. Haworth perspective formula of sucrose with carbon atom numbering.

6-*O*-myristoyl sucrose were shown to differ by a factor of 1.4, and the CMC of 6-*O*-lauroyl sucrose was approximately 2.5 times greater than that of 1′-*O*-lauroyl sucrose (Polat and Linhardt, 2001). The stability of sucrose alkanoates under acidic and basic conditions at elevated temperatures has also been shown to differ among regioisomers. Under acidic conditions (pH 2-4) the regioisomers of sucrose laurate exhibited stability in the order 3′ > 1′ > 6′ > 4′ > 6 > 2 > 4, while under basic conditions (pH 10-12) sucrose laurate was overall more unstable and the different regioisomer stability order 6′ > 1′ > 6 > 4 > 4′ > 2 > 3′ was observed (Okumura et al., 2011). In both cases, the stability of the 3-regioisomer was unknown.

Regioisomer interconversions by acyl migration have also been reported. The acyl migration is promoted by solvent media with acid-base properties, and the general direction of the acyl migration is towards the primary hydroxy positions 6, 1′ and 6′, with migration to the 6-regioisomer favoured (Ritthitham et al., 2009a; Thevenet et al., 1999). However, over shorter time spans a more rapid migration from the 2- to the 3-position has been observed (Molinier et al., 2003). This supports the regioisomers exhibiting differences in physicochemical properties, stability and reactivity.

1.1.2 Antimicrobial properties

Sucrose alkanoates and laurate esters of carbohydrates have been shown to possess antimicrobial properties. For instance, sucrose myristate and sucrose palmitate inhibit spore development in some thermophilic gram-positive bacteria (Piao et al., 2006), and fructose laurate and galactose laurate inhibiting growth of *Streptococcus mutans* (Watanabe et al., 2000).

Sucrose laurate has been shown to inhibit the growth of bacteria across a range of families. Kato and Amira (1971) reported that sucrose laurate at 100 mg/L inhibited growth of *E. coli* while addition of higher concentrations immediately stopped the bacterial growth, and the effect was shown to be greater than that of lauric acid. The ester has also been shown to inhibit the growth of *Bacillus cereus* (Habulin et al., 2008) and *Staphylococcus aureus* (Kato et al., 1987; Monk et al., 1996), and Monk et al. (1996) also reported sucrose laurate as bacteriostatic against *Listeria monocytogenes* at 100 mg/L, and bactericidal at higher concentrations (400 mg/L). The regioisomer 6-*O*-lauroyl sucrose has been reported to inhibit growth of *Bacillus* sp. CECT 40 and *Lactobacillus plantarum* (Ferrer et al., 2005), and also to inhibit growth of *Streptococcus sobrinus* at 100 mg/L (Devulapalle et al., 2004). These antimicrobial effects make sucrose laurate very promising as an ingredient in a variety of food and personal care products, and further investigation of the antimicrobial properties of sucrose laurate regioisomers is warranted.

1.2 Biocatalysis in organic solvents

The use of biocatalysts in organic synthesis offers several general advantages compared to conventional chemical processes. Enzymes are efficient catalysts that function over moderate ranges of pH and temperature and have low environmental impact, as they are completely degradable. They catalyse a broad spectrum of reactions, as there is an equivalent enzyme-catalysed process for almost every type of organic reaction, and they accept a wide range of substrates. Enzymes exhibit generally high selectivity of three major types: chemoselectivity, regioselectivity and stereoselectivity. Chemoselectivity is the preferential reaction at a specific functional group in a substrate molecule, such as a hydroxy, a carbonyl or an amino group. Regioselectivity describes reaction at one molecular position predominating over other positions, including differentiation of reaction sites that are identical functional groups, such as for polyhydroxy compounds like carbohydrates. Stereoselectivity is the preferential formation of, or reaction with, one particular stereoisomer, whether a diastereomer or an enantiomer (Faber, 2004; Nic et al., 2012). Regio- and stereoselectivity is determined by the enzyme secondary, tertiary and quaternary structure, particularly in the active site and its vicinity. Steric and electronic effects from amino acid residues in the crucial area direct the orientation, positioning and conformation of substrate molecules and the enzyme, resulting in enzyme-substrate complexes of varying energetic favourability and different rates of product formation.

Application of biocatalysts in organic solvents provides some additional advantages. In non-aqueous media, water-dependent side-reactions are suppressed, resulting in fewer by-products, and improved solubility of non-polar substrates leads to faster reaction rates. Organic solvents also increase the rigidity of enzymes through impeded hydrogen-bond formation, which can increase the control of substrate specificity and catalytic selectivity (Fitzpatrick and Klibanov, 1991). Overall, the solubility of substrates and products, the propensity for side-reactions and the activity and selectivity of enzymes can be influenced through using different solvents or solvent mixtures with varying physicochemical properties, such as dipole moment (μ), relative static permittivity (ε_r), boiling point (T_B), partition coefficient (logP), acid-base properties and water activity (a_w), a strategy known as solvent or medium engineering. The enzymes known as hydrolases have seen the most research and application as biocatalysts, making the most important advantage of using organic solvents the possibility of shifting thermodynamic equilibria to favour synthesis over hydrolysis, enabling the hydrolase-catalysed formation of compounds such as esters, polyesters, lactones, amides and peptides.

1.2.1 Hydrolase ester synthesis

Ester synthesis in organic solvents can be catalysed effectively by lipases or proteases. These enzymes do not need sensitive cofactors, and a large number of enzymes with a broad range of fairly relaxed substrate specificities, is available. Lipases catalyse ester synthesis either through the condensation reaction between an alcohol and a carboxylic acid or through the acyl transfer between an ester and an alcohol, while proteases only catalyse acyl transfer reactions.

The catalytic mechanism of the hydrolytic enzymes is very similar to the chemical base-catalysed hydrolysis or ester synthesis. The mechanism which has been elucidated in detail is that of the serine-hydrolases, such as most microbial lipases and the subtilisin family of

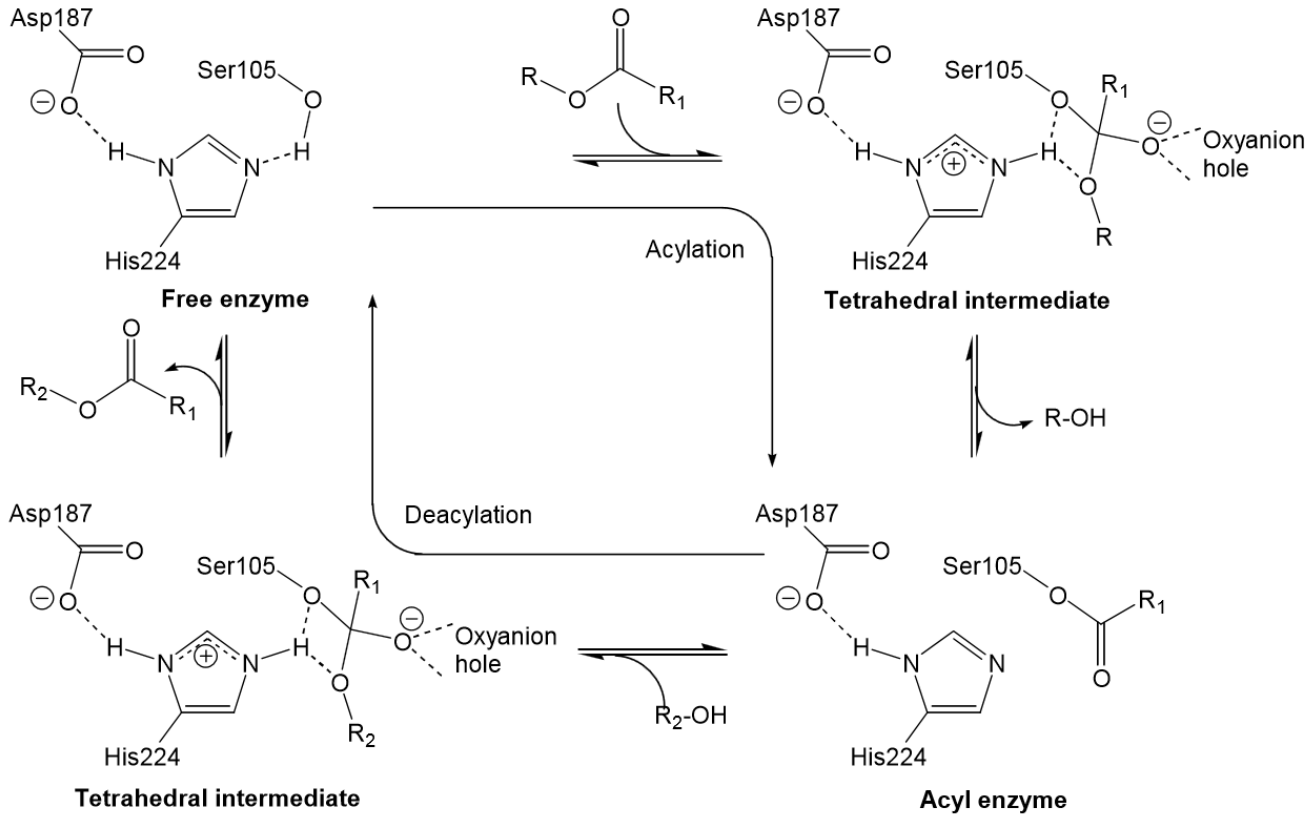


Figure 1-2. The catalytic mechanism of the serine-hydrolases (Ottoson, 2001).
R-O-(CO)-R_1 is the acyl donor, R-OH is the leaving group, and R_2-OH is the acyl acceptor.

proteases (see fig. 1-2). In the catalytic site, a serine residue is part of a hydrogen-bonding network referred to as the *catalytic triad* together with one residue each of aspartate and histidine. The spatial arrangement of these residues decrease the acid-ionisation constant (pK_a) of the serine hydroxy group, enabling the nucleophilic attack on the acyl donor carbonyl group, leading to a tetrahedral intermediate stabilised by coordination to the so-called *oxyanion hole*, and the acyl-enzyme intermediate forms by liberating the leaving group. The incoming acyl acceptor hydroxy group coordinates to the histidine residue, enabling the nucleophilic attack on the acyl enzyme, leading to the regeneration of the enzyme and release of the ester product (Faber, 2004; Ottoson, 2001).

An enzyme applied to the synthesis of carbohydrate esters not belonging to the serine-hydrolases is thermolysin, a thermostable neutral metalloprotease from *Bacillus thermoproteolyticus* characterised by the presence of a zinc cation in the active site. A mechanism for transesterification has been proposed by Perez-Victoria and Morales (2006) based on the accepted hydrolytic mechanism for thermolysin (see fig. 1-3). As the acyl acceptor coordinates to the active site zinc cation, a glutamate residue acts as a general base catalyst. This facilitates the acyl acceptor nucleophilic attack on the acyl donor carbonyl coordinated to a histidine residue, which hydrogen-bond stabilises the tetrahedral intermediate, followed by the liberation of the leaving group and formation of the ester product.

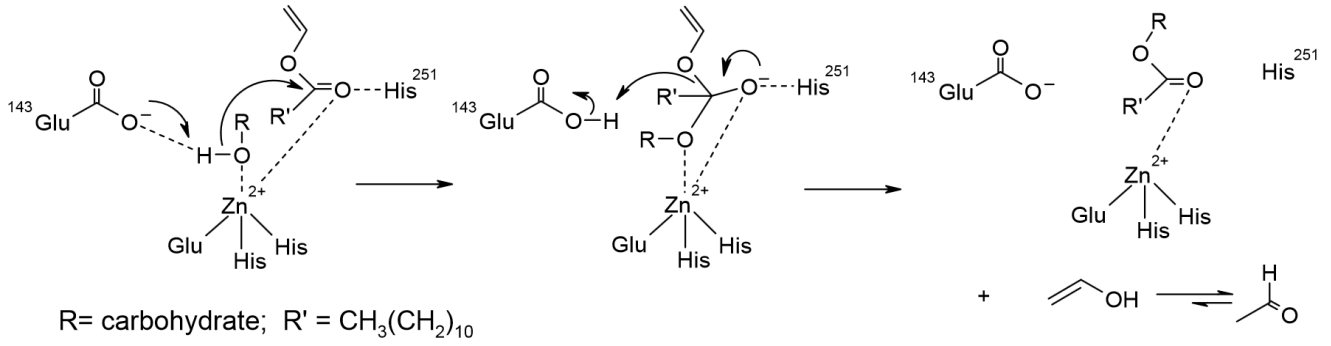


Figure 1-3. Proposed catalytic mechanism for thermolysin (Perez-Victoria and Morales, 2006). R′-(CO)-O-CH=CH_2 is the vinyl ester acyl donor and R-OH is the acyl acceptor.

The reversibility of the ester synthesis reaction can be a challenge as the generally limited substrate concentrations lead to unfavourable reaction equilibria, and with carboxylic acid substrates the formation of water allows for the competing ester hydrolysis and can also lead to substrate solubility issues. Acyl transfer reactions avoid the formation of water, however, controlling the relative nucleophilicity of the acyl acceptor and the leaving group of the acyl donor is still important in order to shift the reaction equilibrium towards the desired product. Esters with electron-withdrawing substituents, such as 2-haloethyl, cyanoethyl, and trihaloethyl esters, give rise to weakly nucleophilic alcohols that make good leaving groups, and using these *activated esters* results in increased synthesis reaction rates. Even more favourable synthesis conditions can be achieved by using enol esters, vinyl or isopropenyl esters that liberate enol leaving groups which tautomerise to the corresponding aldehyde or ketone, leading to virtually irreversible reactions with rapid reaction rates and generally high selectivity.

1.2.2 Biocatalytic synthesis of sucrose alkanoates

Syntheses of sucrose alkanoates catalysed by lipases and proteases in various organic solvents have been reported to exhibit varying degrees of regioselectivity towards several substituent positions (see table 1-1). The most commonly favoured positions are the primary hydroxys 6- and 1′-OH and the secondary hydroxy 2-OH (Shi et al., 2011).

Although the chosen biocatalyst largely determines activity and selectivity, choice of solvent affects several reaction properties, as outlined above (see section 1.2). Polar, aprotic solvents, including *N*,*N*-dimethylformamide (DMF) and dimethyl sulfoxide (DMSO), are excellent solvents for carbohydrates, but can act as protein denaturants, affecting enzyme stability and activity. Alkaline proteases have shown high stability in such solvents, in some cases higher stability in pure DMF than in water (Ogino and Ishikawa, 2001), and improved activity and selectivity with addition of DMSO to the reaction medium have been reported (Almarsson and Klibanov, 1996; Ritthitham et al., 2009b). Lipases have shown sensitivity to high-polarity solvents, losing activity and exhibiting complete inactivation with DMSO concentrations from 20-30 % and above (Degn and Zimmermann, 2001; Reyes-Duarte et al., 2005), and are thus more commonly employed in less polar solvents, such as pyridine and 2-methyl-2-butanol.

Nuclear-magnetic-resonance (NMR) studies and computer modelling of sucrose conformations in polar solvents have shown the chemical reactivity order of the sucrose substituents to be 2-OH > 1′-OH > 3′-OH due to intramolecular conformation-stabilising hydrogen-bonds between 2-OH and either of 1′-OH or 3′-OH (Christofides and Davies, 1985; Lichtenthaler et al., 1995). As outlined in section 1.1.1, sucrose alkanoates are susceptible to acyl migration, and enzymes have also been reported to catalyse acyl migration, such as the alkaline protease AL-89 promoting migration from the 2- to the 3-position (Ritthitham et al., 2009b). Thus, several factors affect the ester formation regioselectivity, and the sucrose alkanoate regioisomer distribution will result from a complex interplay between enzyme selectivity and catalytic activity, chemical reactivity and acyl migration (Queneau et al., 2004; Thevenet et al., 1999).

Table 1-1. Enzymatic synthesis of sucrose esters in organic solvents.
All reactions used activated or enol ester as acyl donor. Adapted from Shi et al. (2011).

Enzyme	Solvent	Time (h)	Products	Yield*	Ref
Thermolysin	DMSO	24	2-*O*-lauroyl sucrose	44 %	[a]
	DMF	24	2-*O*-lauroyl sucrose	52 %	
Thermolysin	DMSO	24	2-*O*-lauroyl sucrose	90 %	[b]
B. pseudofirmus AL-89 protease	DMF – DMSO (1:1, *v*/*v*)	13	2-*O*-stearoyl sucrose 3-*O*-stearoyl sucrose 3′-*O*-stearoyl sucrose	53 % 26 % 4 %	[c]
B. pseudofirmus AL-89 protease	DMF – DMSO (1:1, *v*/*v*)	24	2-*O*-lauroyl sucrose 1′-*O*-lauroyl sucrose 6-*O*-lauroyl sucrose	57 %	[d]
	pyridine – DMF (1:1, *v*/*v*)	24	2-*O*-lauroyl sucrose 1′-*O*-lauroyl sucrose 6-*O*-lauroyl sucrose	53 %	
Subtilisin A	DMF	24	1′-*O*-lauroyl sucrose	-	
	pyridine – DMF (1:1, *v*/*v*)	24	1′-*O*-lauroyl sucrose	60-75 %	
Proteinase N	DMF	120	1′-*O*-lauroyl sucrose	53 %	[e]
B. subtilis protease	DMF	168	1′-*O*-vinyladipoyl sucrose	>90 %	[f]
Subtilisin Carlsberg	DMF	60	1′-*O*-butyryl sucrose	51 %	[g]
Subtilisin Carlsberg	pyridine	24	1′-*O*-lauroyl sucrose	-	[h]
H. lanuginosa lipase w/ Celite	DMSO	24	2-*O*-lauroyl sucrose	46 %	
T. lanuginosus lipase	2-methyl-2-butanol – DMSO (4:1, *v*/*v*)	24	6-*O*-lauroyl sucrose	-	[i]
H. lanuginosa lipase w/ Celite	2-methyl-2-butanol	10	6-*O*-lauroyl sucrose	25 %	[j]
	2-methyl-2-butanol – DMSO (4:1, *v*/*v*)	24	6-*O*-lauroyl sucrose	73 %	
	2-methyl-2-butanol – DMSO (4:1, *v*/*v*)	48	6-*O*-palmitoyl sucrose	80 %	
Bioenzyme 240	DMF	120	1′-*O*-butyryl sucrose	100 %	[k]

***** The yields are not suitable for direct comparisons, as they are reported differently across the sources. [a] (Pedersen et al., 2002a); [b] (Pedersen et al., 2002b); [c] (Ritthitham et al., 2009b); [d] (Pedersen et al., 2003); [e] (Potier et al., 2001); [f] (Borges and Balaban, 2007); [g] (Riva et al., 1988); [h] (Plou et al., 1999); [i] (Ferrer et al., 2005); [j] (Ferrer et al., 1999); [k] (Patil et al., 1991)

1.3 Reversed-phase high-performance liquid chromatography

For some decades, high-performance liquid chromatography (HPLC) has been the analytical technique of choice for product or sample separation and detection in areas such as pharmacology, biotechnology and applied sciences. HPLC is today a versatile analysis technology suitable for application to a wide range of chemical compounds and analytical fields through adaptability of several components, such as columns of different geometries and various stationary phases, selection and mixing of different mobile phase solvents and a wide range of detection methods (Swartz, 2010).

Reversed-phase (RP-) HPLC operates on the principles of hydrophobic interactions and use nonpolar stationary phases and aqueous, moderately polar mobile phases, resulting in longer retention times for less polar analytes, while polar molecules are eluted more rapidly. The most popular columns for RP-HPLC are based on alkyl substituents covalently bonded to silica support materials, such as octadecyl- or octyl-groups (C18 and C8, respectively), while widely-used eluents are aqueous mixtures of acetonitrile or methanol. Mobile phase modifiers are often added in low concentrations to enhance the efficiency of the separation, for instance buffering agents to control pH and neutralise charge in analytes or on the stationary phase surface. Trifluoroacetic acid (TFA) is a common modifier in analyses involving carboxylic acids, as it is a very strong acid and soluble in a wide range of solvents (Sadek, 2004). RP-HPLC has been shown to be suitable and effective for the analysis of sucrose alkanoates (see section 1.3.3).

1.3.1 Chromatographic parameters and resolution of analytes

Chromatographic analysis results in an output of several analytical parameters that describe the characteristics of the sample analytes as recorded in a chromatogram. Retention time (t_R) indicates peak positions and distribution of analytes throughout the run time, peak height (h) and peak area (A) reflect amounts of analyte, while the asymmetry factor (A_s) describes deviation from the ideal Gaussian shape for each peak, expressed as the ratio

$$A_s = b/a \tag{1-1}$$

where a and b are the widths of each peak-half at 1/10 of the peak height (see fig. 1-4) (Sadek, 2004). The peak width at baseline is defined as $w_b = 4\sigma$, that is, double the width at the Gaussian inflection points. Two parameters also provide information on the chromatographic run as a whole: the dead time (t_0), which is the elution time of a completely unretained compound; and the analysis time, which depends on the last eluting analyte.

To determine the extent of analyte separation from the chromatographic peaks, visual appraisal of a chromatogram has severe limitations. The realisation of a chromatogram on paper or on screen is strongly dependent on the scales and ranges of the time and signal axes and changes in the level of magnification can make a pair of closely eluted peaks appear more or less overlapping to such a degree that objective assessment of the separation beyond the appearance of two peaks becomes essentially impossible.

Resolution (R_s) is a metric that characterises the extent of overlap between pairs of peaks. It is defined as the ratio of the distance between peak apexes to the average peak width, expressed mathematically as

$$R_s = \frac{t_2 - t_1}{\frac{1}{2}(w_1 + w_2)} = \frac{2(t_2 - t_1)}{w_1 + w_2} \quad (1\text{-}2)$$

where t_1 and t_2 are the retention times and w_1 and w_2 are the peak widths at baseline for any pair of successive peaks with $t_1 < t_2$ (Tijssen, 1998). Satisfactory resolution, that is, minimal peak overlap, corresponds to the range 1-1.5, but the degree of peak overlap for a given R_s-value depends on the peak size ratio. The rule of thumb is that optimised resolution, also called baseline separation, is achieved for $R_s \geq 1.5$. However, in most practical situations $R_s \geq 1.0$ is considered the threshold for analytical quantification (Tijssen, 1998).

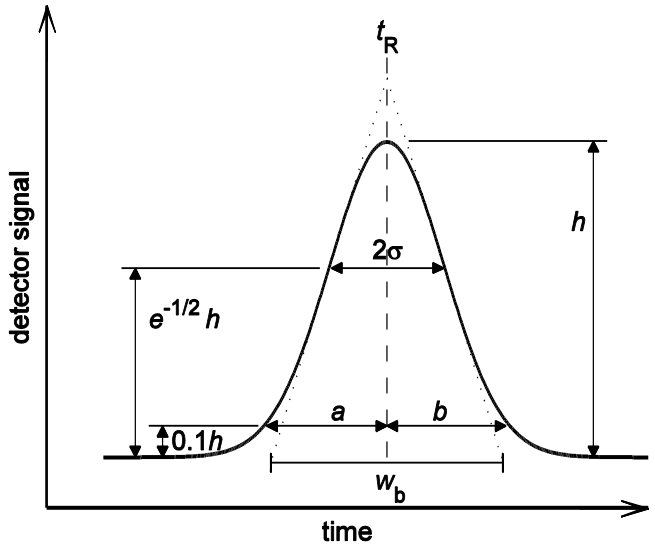


Figure 1-4. Chromatographic peak with ideal shape as a Gauss curve.
Analytical parameters dependent on peak size and peak shape are indicated.

1.3.2 Charged aerosol detection

The detector most commonly used with HPLC today is the UV-visible absorbance detector. This detector responds to light absorption at wavelengths in the range from 190 to 600 nm and is thus dependent on the presence of chromophores in the analysed compounds. This makes it less suited for RP-HPLC analysis of carbohydrates and their fatty acid ester derivatives, as these analytes are weakly light absorbing and absorb in ranges overlapping with typical mobile phase solvents, such as acetonitrile.

The charged aerosol detector (CAD, also known as a Corona CAD) is based on the trapping of charged analyte particles following post-column changes to the mobile phase, and thus not directly dependent on analyte physicochemical properties. The detector functions by nebulising the eluant with a nitrogen gas flow and transporting the resulting aerosol through a drying tube, where volatile components and solvents are evaporated, producing analyte particles. The dried particle stream is diffusionally charged by a secondary stream of nitrogen which has passed a high-voltage, platinum corona wire, and the resulting charged particles are collected and the

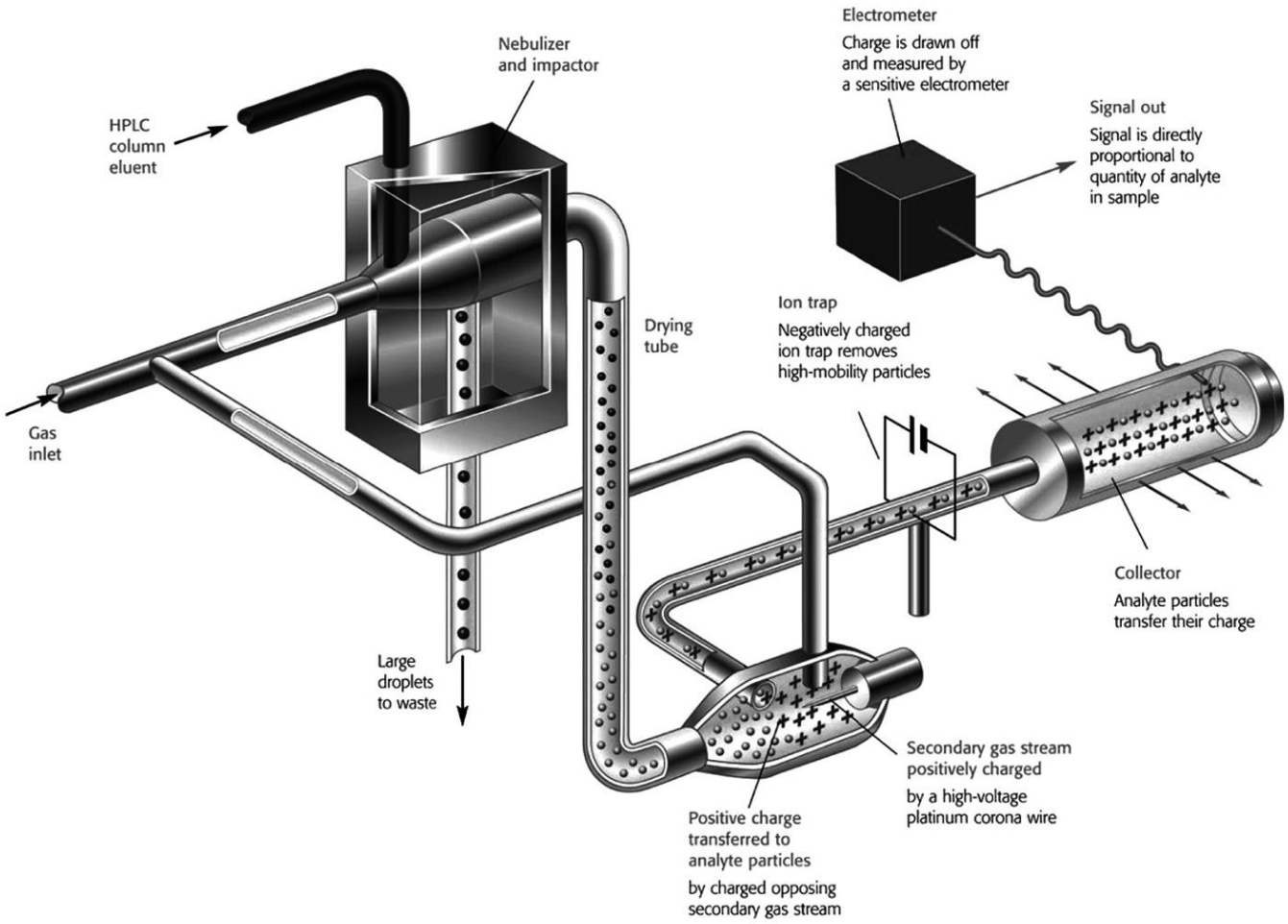


Figure 1-5. Simplified schematic of a corona charged aerosol detector (Swartz, 2010).

charge flux measured by a highly sensitive electrometer, generating a signal in direct proportion to the quantity of analyte present (see fig. 1-5) (Swartz, 2010; Vehovec and Obreza, 2010).

Touted as a 'universal' detector when developed, CAD is highly sensitive, provides a consistent response, and has a broad dynamic range. Some limitations that apply to this detection method are sensitivity to analyte volatility and to organic solvent eluent concentration. The CAD is suitable for analysis of non-volatile compounds, however, this is of course relative to the operating conditions of the detector, and especially the temperatures of the nebuliser and drying tube, where examples are the ESA Corona CAD operating at ambient temperature, and the Dionex Corona Veo CAD which has a nebuliser thermostat with settings of 35 or 50 °C. The effect of eluent composition stems from the influence of organic solvents on the nebulisation and droplet evaporation processes in the detector, which can lead to changes by factors of up to 4-5 in the response of individual analytes due to variations in the transport efficiency of particles within the detector (Hutchinson et al., 2010). This causes some baseline drift and varying detector response for analytes eluted at different eluent compositions when used with gradient elution.

CAD exhibits characteristic non-linear mass-response relations which follow power functions expressed by the equation

$$A = aM^b \tag{1-3}$$

where *A* is the peak area response from the detector, *M* is the mass injected and *a* and *b* are constants characteristic for specific analytes and chromatographic conditions. The response has been found to be lower for particles above 10 nm in diameter (Dixon and Peterson, 2002) which explains the exhibited exponents, *b*, with values typically below 1.0 (Hutchinson et al., 2010; Vervoort et al., 2008).

Equation 1-3 can be transformed into

$$\log A = b \log M + \log a \tag{1-4}$$

and a log-log plot of the peak area versus analyte mass will result in a linear calibration curve. Following log-log transformation, linear regression can be used to determine the coefficient, *b*, and the intersect term, log *a*. To determine analyte mass from peak area response, eqs. 1-3 and 1-4 are combined to the expression

$$M = \left(\frac{A}{e^{\log a}}\right)^{\frac{1}{b}} \tag{1-5}$$

where the values from the regression equation are inserted for calculation.

Other detection methods that have been used with sucrose alkanoates are evaporative light-scattering detection (ELSD), refractive index detection (RID) and coupled mass spectrometry (LC-MS), but CAD offers clear advantages in terms of sensitivity, dynamic range, and ease of use (Swartz, 2010). Specifically for carbohydrates, fatty acids and lipids, CAD has been shown to be a suitable high-sensitivity detection method (Fox et al., 2013; Tamura et al., 2010; Vehovec and Obreza, 2010).

1.3.3 RP-HPLC analysis of sucrose alkanoates

Sucrose alkanoates have been analysed by RP-HPLC for at least two decades (Torres et al., 1990), proving this chromatographic method well suited for these analytes. However, the main focus has mostly been on the separation of alkanoates of different chain lengths or degrees of esterification, and although the presence of sucrose alkanoate regioisomers has been indicated by peak groups identified as the same compound, little effort has been made to further resolve these (Ferrer et al., 2005; Moh et al., 2000; Wang et al., 2007b).

There have been reports claiming to achieve separation of either different alkanoates or alkanoate regioisomers without showing any chromatographic results and describing only the chromatographic conditions employed (Huang et al., 2010; Pedersen et al., 2003; Reyes-Duarte et al., 2005). This provides no references for attempts to replicate the results and no starting point for further investigation towards improvement of the analysis methods. Other reports of sucrose alkanoate regioisomer separation include chromatograms demonstrating the achieved separation, but without providing or discussing the achieved resolutions (Christiansen et al., 2011; Cruces et al., 2001; Okumura et al., 2011; Wang et al., 2006).

More successful separations of sucrose alkanoate regioisomers using RP-HPLC with C18-columns and acetonitrile-in-water eluents have been reported. Perez-Victoria et al. (2007) reported the separation of all eight regioisomers of sucrose caprate, sucrose laurate and sucrose palmitate by isocratic elutions, without presenting resolution data. Ritthitham et al. (2009a) reported using preparative RP-HPLC to achieve separation of seven regioisomers of sucrose caprate by gradient elution, but without baseline separation for a group of three regioisomers and no discussion of resolutions. It was also reported that improved separation of the sucrose caprate regioisomers was obtained by decreasing the eluent acetonitrile concentration from the initial concentration, followed by an increasing gradient, representing a rather untraditional combination of reverse and multistep linear gradient profiles (Sadek, 2004).

1.4 Strategy of experimentation and analysis of results

Experimentation and analysis of resulting data are important in scientific research to obtain valid and objective conclusions. Observation of a system or process can lead to theories or hypotheses, but experimentation is necessary to elucidate information about how and why it works and to construct empirical models of the system under investigation.

In general, a system or process can be visualised as a combination of parameters or factors that transforms some input into an output that consists of one or more measurements, or has one or more observable properties, called the response variables (see fig. 1-6). Some of the process variables or factors, x_1 to x_k, will be controllable, while other variables, z_1 to z_q, are uncontrollable and known as nuisance factors (Montgomery, 2012). Experiments are mostly devised to determine the influence of the different factors on the responses of the system or process, and the general approach to planning and conducting experiments for this purpose is called the strategy of experimentation.

Simple strategies of experimentation are the best-guess and one-factor-at-a-time (OFAT) approaches. The direct strategy is the best-guess or *ad hoc* approach, where arbitrary combinations of factors are chosen and tested, and results are assessed based on knowledge and experience. Although this strategy can work in practice, it is very hit-or-miss, can be

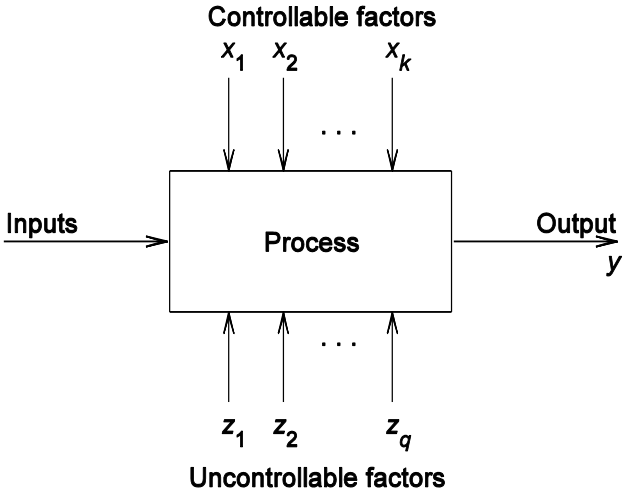


Figure 1-6. General model schematic of a process or system (Montgomery, 2012).

time consuming and provides little certainty that the best result has been found. The traditional scientific strategy is the OFAT approach. This method is based on the successive variation of each investigated factor over its range while all other factors are held constant. Analysing the responses provides results for the influence or effect of each factor in isolation. This is the major disadvantage of the OFAT approach, as it assumes the factors to be independent, and cannot account for the possibility of interactions between them and how these influence the responses. In addition, this approach cannot distinguish random variations from the true factor effects and is limited to the results of the experiments actually performed (Esbensen et al., 2002).

Design of experiments (DOE) is a strategy based on statistical theory which makes it possible to investigate all factors together and still draw appropriate conclusions about individual and interaction effects independently from each other. DOE can be applied for a variety of purposes, such as characterisation of a process, optimisation of a process, product design or product formulation. Advantages associated with DOE are efficiency and focus, as more information is gained from fewer experiments, only the needed information is collected, and the necessary number of experiments to obtain the desired information is known. In addition, the effects of individual factors and their interactions can be studied simultaneously, true effects can be distinguished from random variation, conclusions can be made about the statistical significance of observed effects and the impact of experimental errors on the results is at a minimum (Esbensen et al., 2002; Montgomery, 2012).

The basic principles of DOE are randomisation, replication and blocking. The statistical foundation of the strategy imposes randomisation of the order of the individual experimental runs to average out the effects of nuisance factors and avoid bias in the results, while replication of the individual runs is necessary to obtain an estimate of the experimental error and improves the precision of the average responses at each factor level. Blocking is the grouping of experiments according to an uncontrollable factor, such as raw material batches or time constraints preventing the completion of the full design in one session or as a continuous series, which improves the precision of factor comparison and can reduce or eliminate variability from such nuisance factors (Dean and Voss, 1998).

The practical process of experimentation using DOE can be summarised as: planning, which involves developing a problem statement, choosing response variables, choosing factors and their levels and ranges; choice of experimental design; performing the experiment; statistical or multivariate analysis of data; and forming conclusions and recommendations from the results. Throughout this process it is important to make use of problem-specific, non-statistical knowledge and to recognise the difference between practical and statistical significance to inform choices and decisions. In addition, such experimentation is an iterative process, that is, initial hypotheses are formulated and investigated through experimentation, and on the basis of the results the hypotheses are reformulated, and so on.

The construction of experimental designs mostly makes use of the balanced unity range, ±1, or an equivalent suitable, balanced normalised range, for all design variables. These coded variables allow universal application and adaptability of the design, and, in the statistical analysis, make the estimated effects (or model coefficients) directly comparable and estimated

with the same precision. This facilitates ease of interpretation, making coded variables very effective for determining the relative size of factor effects. In practice, construction of experimental designs and analysis of data are conducted using computer software. Several software packages can be employed for this purpose, such as R (The R Foundation for Statistical Computing), MATLAB (MathWorks) and The Unscrambler (CAMO Software).

1.4.1 Experimental designs

The choice of experimental design depends on decisions about which factors to change, the intervals for this variation and the pattern of the experimental points, all of which should be informed by the objective of the experiment (Esbensen et al., 2002). Depending on the number of factors and whether the objective is, for instance, the screening of significant factors or the optimisation of the process, different experimental designs will be suitable.

The basic experimental design for investigating the joint effect of several factors on a response is called a factorial design. This design incorporates all possible combinations of the levels of the factors in each complete trial or replicate and allows for the study of the main effects of all the factors and interaction effects between them. General factorial designs can be constructed for any number of factors at individual numbers of levels. The most widely employed factorial design is called a 2^k-factorial design, as it investigates k factors, each at two levels, in order to estimate main effects and interactions and their significance. This design is often used for variable *screening*, that is, to determine the importance of each factor and whether the number of factors studied in more detail can be the reduced.

The addition of centre points to a factorial design allows for the detection of curvature in the relationship between factors and responses. The centre points are performed at the mean level for all factors and do not affect the usual effect estimates in the design, however they allow the estimation of a quadratic effect and its significance, and an independent estimate of error is obtained from replicates of this additional experimental run. This error estimate increases the efficiency of the experimental design by reducing the need for replicating the whole design, while obtaining very similar results (Esbensen et al., 2002; Montgomery, 2012).

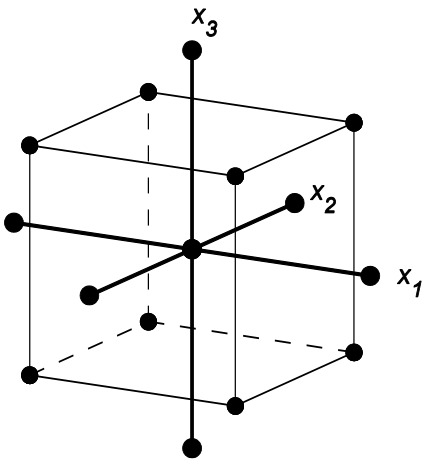


Figure 1-7. Circumscribed central composite (CCC) design for three factors.

Central composite designs (CCD) are more advanced experimental designs which investigate factors at more than two levels to analyse the results with a more complex model and predict the parameters for the optimal response. These *optimisation designs* incorporate centre points and are used extensively for second-order response surface modelling. The general CCD for k factors extends a 2^k-factorial design with $2k$ axial runs and centre points, and investigates each factor at five levels in various combinations over a total of $(2^k + 2k + 1)$ individual runs. The standard CCD for 2 factors can be visualised as a circle circumscribed on the factorial square, while a 3-factor CCD equates to a sphere circumscribed on the factorial cube (see fig. 1-7), which is why they are also called circumscribed central composite (CCC) designs. These designs possess a property called rotatability, meaning that each experimental run contributes equally to the total information, and the resulting model will have equal precision of estimation in all directions from the centre of the design. Normalised to the factorials, the distance of the axial points from the centre is $\sqrt{k}$ in such a design (Montgomery, 2012).

In many situations there might be practical limitations to the experimental ranges of the investigated factors, leading to the cuboidal region of operability coinciding with the region of interest. A variation of the CCD useful for such cases is the face-centred central composite design or face-centred cube (FCC), in which the axial runs are constrained to the factorial ranges, and each factor is investigated at three levels. For a 3-factor design, the axial points lie on the faces of the factorial cube (see fig. 1-8). A disadvantage of FCCs is their lack of rotatability, as the axial runs contribute slightly less information than the factorials, and the resulting model will not have as high precision of estimation as for a CCD, with the difference becoming more pronounced towards the edges of the design space (Montgomery, 2012).

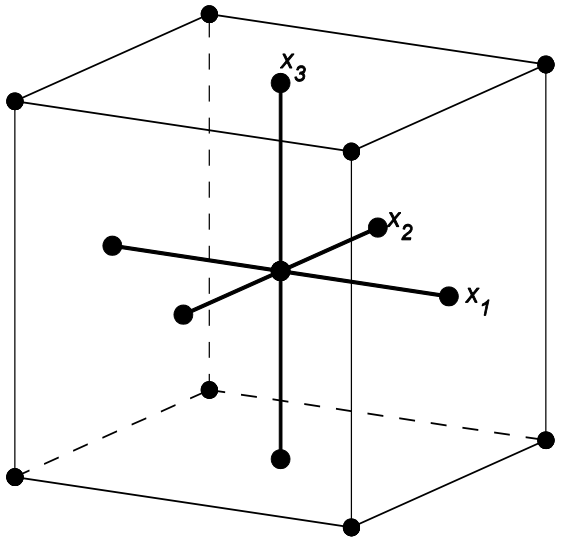


Figure 1-8. Face-centred cube (FCC) design for three factors.

1.4.2 Statistical analysis

Analysis of DOE data strives to determine the relationship between the factors and responses in the studied system. The effect of a factor on a response is the change in the response produced by a change in the level of the factor. Taking a 2^2-factorial design as an example, the main effect of each factor is estimated as the difference between the average responses at the high and low levels of that factor, while the interaction between the factors is estimated as the average difference between the effects of one factor at the high and low levels of the other factor. The coefficients of the regression model for this factorial design are half the effect estimates, while the intercept term is equal to the grand average of all the responses.

Beyond simple designs with few factors, the relationship between the factors and responses in a system is better approximated using multiple linear regression (MLR) modelling, and in any case, the use of such statistical methods is necessary in order to determine the adequacy or validity of the model objectively. MLR is a special case of analysis of variance (ANOVA), a statistical method that partitions the total variability of a studied system into its component parts, representing the factors, as well as a part due to error. The application of ANOVA to an MLR model makes it possible to determine the statistical importance of the studied design variables in terms of influence on the response, and to assess how well the model approximates their true relationship through different techniques, known as model adequacy checking.

The mathematical model of a system can be expressed as the response observations, y, being a function of the design variables, x, the model parameters, β, and the error terms, ϵ, written in matrix notation as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1-6}$$

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

for n experimental runs, k variables, and k + 1 model parameters, including an intercept parameter, β_0 (Montgomery, 2012). Least squares estimation of the model parameters minimises the sum of squares of the errors, $\boldsymbol{\epsilon}$. The least squares estimator of $\boldsymbol{\beta}$ is determined as

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y} \tag{1-7}$$

and from eqs. 1-6 and 1-7, the fitted regression model in matrix notation becomes

$$\hat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}} \tag{1-8}$$

or as expressed in scalar notation

$$\hat{y}_i = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \quad i = 1, 2, \ldots, n \tag{1-9}$$

MLR can estimate model parameters for polynomial models incorporating variable interactions (e.g. x_1x_2) and variable exponentials (e.g. x_1^2) by representing these complex terms as

additional design variables in the above expressions (eqs. 1-6 to 1-9). The maximum number of model parameters that can be estimated, including the intercept term, is equal to the number of independent, or unique, experimental runs in the design.

When applying ANOVA to a regression model, the factors investigated and the terms for which model parameters are estimated represent the variance components into which the total variability is partitioned. Statistical significance is determined by considering each ratio of component variance to the residual error variance against the appropriate probability density function (*F*-distribution) resulting in a probability (*p*-value) of that variance being due to random error. If the *p*-value is below the chosen significance level (usually $\alpha = 0.1$ or 0.05), the hypothesis that the variances are equal is rejected, and the effect is considered statistically significant. Thus, the ANOVA results and *p*-values obtained indicate the relative importance of the terms in the system and which terms in the regression model have statistically significant effects on the response.

Models including non-significant terms can be simplified by elimination of the least significant, highest-order term, with the expected result that the reduced model will better model the data. Regression modelling and ANOVA are repeated with the reduced number of terms, resulting in new *p*-values for the model terms. This elimination of model terms should follow the *hierarchy principle*. This states that any main effect, variable exponential or variable interaction of a specific design variable should not be eliminated before a more complex term containing that variable. For polynomial regression models of several design variables, this can lead to the inclusion of non-significant terms in the final model.

Experimental designs which include centre point replicates allow the residual error variance to be further partitioned into components representing a model-independent pure error variance and a model lack-of-fit variance. This provides an additional test of the model suitability by comparing the lack-of-fit variance to the pure error as outlined in the previous paragraph, and if determined significant, the regression does not model the true relationship between the design variables and the response.

MLR and ANOVA are parametric statistical methods based on some assumptions about the variables and the error variance distribution, although the methods have been shown to be quite robust against departure from these assumptions (Montgomery, 2012). Before the statistical analysis it is nevertheless advisable to assess the distribution of the responses using a normal probability plot and the Shapiro-Wilk normality test (Shapiro and Wilk, 1965). This will indicate whether variance-stabilising, mathematical transformation of the data is necessary. If so, a suitable transformation parameter can be determined empirically or estimated using the Box-Cox method (Montgomery, 2012).

The analysis of DOE data starts with MLR-modelling and ANOVA of the most complex and extensive regression model that can be estimated for the experimental design, and is considered suitable based on problem-specific knowledge. The ANOVA significance results are considered, and if possible, the model is iteratively reduced until only significant and hierarchically required terms remain. To ascertain whether this final model describes the system without issue, it is subjected to model adequacy checking. This consists of considering the

significance of the model as a whole and of the lack-of-fit (if available), assessing the size of the coefficients of determination (R^2 and $R^2_{adj.}$), and appraising the distribution and variance of the model residuals by plotting them against the order of the experimental runs and in a normal probability plot. The root-mean-square-error (RMSE) is also a measure of the average modelling error calculated from the residuals.

Having established the model as adequate, the relative sizes of the design variable effects and interactions can be compared and the significance of the model terms considered. These results provide information on the factor-response relationships in the system. The model can also be used for plotting of response surfaces and contour plots to visualise the factor-response relationships, to predict results or to optimise the studied system. Finally, it will often be appropriate to construct a regression model using the uncoded, actual design variables for direct prediction of the response from the experimental parameters.

1.4.3 Multivariate data analysis

A process or system with output consisting of several response variables, y_1 to y_m, will require separate regression modelling and statistical analysis for each response using MLR and ANOVA, before the resulting response surfaces or contour plots can be used to identify regions of optimality. This might be applicable for two or three responses, but as the number of responses increases to four and above, this process can become highly cumbersome.

Multivariate data analysis offers alternatives that allow for analysis and modelling of the data structures in several design and response variables simultaneously. The method most appropriate for the analysis of DOE data is best known as partial least-squares (PLS) regression, but has also been called projection on latent structures. The last name comes from this being a projection method that, potentially, reveals hidden structures in the data.

In order to partition the total variability of the system into a structural part due to factor effects and a noise part due to random error, PLS sequentially projects the design and response data to new orthogonal axes in each variable space along the multidimensional directions of maximum variance in the response data. The first of these new dimensions, called PLS-components, is determined as the direction in X-space that accounts for as much of the variability in the y-data as possible, while each succeeding PLS-component has the highest variance possible under the constraint that it is orthogonal to the preceding components. This process can also be described as sequential least-squares estimation of the X-Y-relationship, one dimension at the time, by way of iterating a set of interdependent, criss-cross X-/Y-space regressions (Esbensen et al., 2002).

The general underlying mathematical relations of PLS is

$$\begin{aligned} \mathbf{X} &= \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E} \\ \mathbf{Y} &= \mathbf{U}\mathbf{Q}^{\mathrm{T}} + \mathbf{F} \\ \hat{\mathbf{Y}} &= \mathbf{X}\hat{\mathbf{B}} = \mathbf{X}\mathbf{W}(\mathbf{P}^{\mathrm{T}}\mathbf{W})^{-1}\mathbf{Q}^{\mathrm{T}} \end{aligned} \tag{1-10}$$

where **X** is an *n*-by-*k* matrix of the design variables; **Y** is an *n*-by-*m* matrix of the response variables; **T** and **U** are *n*-by-*a* matrices of projection *scores* for **X** and **Y**, respectively;

P and **Q** are, respectively, *k*-by-*a* and *m*-by-*a* matrices of *loadings* for **X** and **Y**; **E** and **F** are matrices of the error terms; **W** is a *k*-by-*a* orthogonal matrix of *loading weights*; and $\hat{\mathbf{B}}$ is an *n*-by-*a* matrix of regression coefficients, for *n* experimental runs and *a* PLS-components. The matrix decompositions of **X** and **Y** are performed so as to maximise the covariance of **T** and **U**, reflecting the description of the PLS-progress in the above paragraph.

In the resulting PLS-model, the scores are the projected values of each experimental run on the PLS-components, while the loadings describe the relations between the raw variable data and the respective scores. The loading weights are vectors describing the actual directions of the PLS-components in the X-space. All of these values describe aspects of the relationships within each of the x- and y-data, and between them.

An important aspect of PLS regression is choosing the number of PLS-components to include in the final model. The primary measure considered in this context is the cumulative explained variance of the PLS-components for the x-data and y-data, with the explained variance being the ratio of the variance modelled by each component to the total variance in each variable space. Usually the first few PLS-components will explain much of the variance in both **X** and **Y**, and a common guideline is to choose the lowest number of components that bring the explained variance above or close to a threshold, often 95 %, reminiscent of choosing the significance level in statistical methods. In addition, the coefficients of determination and RMSE for the model with different numbers of PLS-components should be considered.

The analysis of DOE data using PLS regression is a corresponding process to statistical regression and analysis. The most extensive and complex regression model for the experimental design again serves as the starting point. Although PLS does not result in statistical *p*-values for the estimated regression parameters, the significance of terms can be estimated using cross-validation coefficient-variance significance testing in conjunction with the relative sizes of the regression coefficients. The model is then iteratively reduced by eliminating the least significant variable term according to the hierarchy principle (see section 1.4.2), until only significant effects remain. Model adequacy checking for PLS regression is performed in the same way as for MLR and ANOVA.

The results from the final regression model will describe the relations between the multidimensional variable spaces, although the modelled relations do not represent the actual response variables, but rather the previously hidden covariance dimensions corresponding to the number of PLS-components chosen. The PLS-model can be utilised for the same purposes as MLR-models, however, plotting of PLS-models involving several design and response variables will often be a challenge, and direct prediction is thus the most usual application.

2 DEVELOPMENT AND OPTIMISATION OF RP-HPLC METHODS

To develop methods for quantitative analysis of sucrose alkanoate regioisomers, a wide range of elution strategies for RP-HPLC was investigated, using both a conventional one-factor-at-a-time approach and design-of-experiments (DOE) methodology. The potential and limitations of methodology transfer for the analysis of sucrose alkanoates of increasing alkanoyl chain lengths were also assessed.

Sucrose caprate was analysed using isocratic elution over a range of eluent acetonitrile concentrations and using complex step-down gradient elutions (Ritthitham et al., 2009a). Sucrose laurate was also analysed using isocratic elution and step-down gradient elutions, as well as with combinations of isocratic elutions and different flow rates, and using step-up gradient elutions. Results from the various elution methods that were analysed and evaluated were regioisomer retention times (t_R, min.), average adjusted retention time ($\bar{t}'_R$, min.), analysis time (min.) and *general resolution deviation for multiple peaks* (RD_m), a sample-independent resolution metric for the overall separation obtained with any elution method (see section 2.2).

The RP-HPLC analyses described below were performed with chromatographic conditions as described in Papers 1 and 2 (Part II).

2.1 Design variables

Design-of-experiments approaches were employed for all elution strategies incorporating more than a single factor of variation or design variable. The variables are summarised in table 2-1, and the variable ranges were determined from preliminary isocratic elutions and range testing experiments.

Based on the results from Ritthitham et al. (2009a), the step-down gradient elutions were defined by four variables, describing variations in the elution parameters in the early stages of the elution and followed by a constant eluent concentration. An initial eluent concentration of acetonitrile, *A* (%), was maintained for a duration, *B* (min), after which the acetonitrile concentration was changed to a new level, *C* (%), which was maintained for a duration, *D* (min). The following constant eluent concentration was set according to the fatty acid chain length (35 % for sucrose caprate, 38 % for sucrose laurate) and maintained for the remainder of the analysis time (see fig. 2-1a). The durations of the concentration transitions were kept constant (0.5 min). The flow rate effects under isocratic conditions were investigated using the two design variables eluent flow rate, *F* (mL/min), and isocratic acetonitrile concentration, *I* (%).

Table 2-1: Design variable symbols, definitions and ranges.

Variable	Description	Sucrose caprate	Sucrose laurate
A (%)	Initial acetonitrile concentration	30.0-40.0	38.0-46.0
B (min)	Duration of *A*	2.0-6.0	0.0-6.0
C (%)	Successive acetonitrile concentration	20.0-40.0	24.0-38.0
D (min)	Duration of *C*	1.0-5.0	2.0-6.0
F (mL/min)	Eluent flow	-	1.0-2.0
I (%)	Isocratic acetonitrile concentration	-	35.0-41.0

2.2 General resolution deviation for multiple peaks

Optimisation of the separation of eight sucrose alkanoate regioisomers required the evaluation of simultaneous changes in seven resolutions, which motivated the development of an aggregate objective function (AOF) to serve as a single metric for the overall separation obtained with different elution methods (García-Álvarez-Coque et al., 2006).

The first AOF developed, denominated *resolution deviation* (RD), was based on the comparison of the two R_s-values describing the separation of the regioisomers 6, 3 and 1′ (as annotated in fig. 2-1) to optimal resolution, as analyses had shown that the indicated regioisomers were the least resolved analytes. The numeric comparison was derived from the root-mean-square deviation (RMSD), which is a measure of the accuracy of a set of values compared to their average, the target value, and expressed as

$$\mathrm{RMSD}(x) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2-1}$$

Where n is the number of observations, and $\bar{x}$ is the sample mean or the target value for comparison. The optimal resolution is obtained when $R_s = 1.5$, which describe baseline separation between chromatographic bands (Tijssen, 1998), and any further increases in R_s would be neither beneficial nor detrimental. Thus, optimal resolution was chosen as the target value, $\bar{x} = R_s = 1.5$, and the comparison was constrained by taking the lesser value of zero and the target value subtracted from resolution, or

$$\min(x_i - 1.5,\quad 0) = \begin{cases} 0 & x_i > 1.5 \\ x_i - 1.5 & x_i \leq 1.5 \end{cases} \tag{2-2}$$

For the two resolutions, $n = 2$, and with the expression scaled to an output range of [0, 1] to facilitate easy assessment (by multiplying with $a = 1/\sqrt{2 \times 1.5^2}$), eqs. 2-1 and 2-2 were combined to define RD as

$$\mathrm{RD}(R_{s1}, R_{s2}) = a \times \sqrt{\min(R_{s1} - 1.5, 0)^2 + \min(R_{s2} - 1.5, 0)^2} \tag{2-3}$$

The RD provided a single measure for evaluating the separation achieved by any elution strategy applied in the current investigation. However, the metric was dependent on a group of three neighbouring peaks providing two resolutions, making the RD only suitable for comparisons of elution strategies for any specific sample, as resolution-values are inherently dependent on peak sizes and thus the relative amounts of analytes (see section 1.3.1).

To improve the applicability, RD was further developed into a generalised sample-independent metric, the *general resolution deviation for multiple peaks* ($\mathrm{RD_m}$), incorporating the separation of any number of analytes and normalising their relative amounts.

The RMSD forming the basis for the AOF already provided for any number of input values, while the normalisation of analyte amounts could be accomplished by estimating peak properties for all peaks from the properties of a single peak. By normalising to the peak with the largest (maximum) area, the function avoided overestimation of the resolutions.

Peak width has been shown to be a function of retention time (Tijssen, 1998), giving the slope of the peak width, $\bar{\omega}$, for the largest peak as

$$\bar{\omega} = \frac{w_b}{4t'_R} \quad \text{for max(A)} \tag{2-4}$$

where w_b is peak width at baseline and t_R' is the adjusted retention time. By estimating peak widths for all peaks from $\bar{\omega}$ and assuming that pairs of peaks with similar properties were closely eluted, eqs. 1-2 and 2-4 were combined to the expression for the estimated resolution

$$R_s = \frac{2(t_2 - t_1)}{w_1 + w_2} = \frac{2\Delta t_R}{4\sigma + 4\sigma} = \frac{\Delta t_R}{4\sigma} = \frac{\Delta t_R}{4\bar{\omega} t'_R} \tag{2-5}$$

The estimated resolutions for pairs of successive peaks, (t_{i-1}, t_i), combined with eqs. 2-1, 2-2 and 2-5 gave the final expression for RD_m

$$RD_m(t_R) = a \times \sqrt{\frac{1}{n-1} \sum_{i=2}^{n} \min\left(\frac{\left(t_i - t_{i\text{-}1}\right)}{4\bar{\omega}(t_i - t_0)} - 1.5, 0\right)^2} \tag{2-6}$$

Where $i \in [2, n]$, as the number of calculated resolutions is always $n - 1$, that is, one less than the number of peaks evaluated, and a is a scaling constant. Scaling by a = 1/1.5 resulted in a range between 0 and 1, $RD_m \in [0, 1]$, where $RD_m \propto 1/\sum R_s$.

Some particular properties of the AOF should be pointed out. RD_m has a decreasing scale and approaches zero as the resolutions approach the target value indicating baseline separation. In addition, the metric only indicates the overall resolution of the analytes, and does not reveal anything about the relative sizes of the component resolutions, as several combinations of R_s-values can result in the same RD_m.

2.3 Separation and identification of sucrose alkanoate regioisomers

RP-HPLC analyses of the sucrose alkanoates resulted in the resolution of the eight regioisomers for sucrose caprate and sucrose laurate, with R_s-values ranging from 1.31 to 6.82 for sucrose caprate and from 1.34 to 7.68 for sucrose laurate (see fig. 2-1). Analyses of control samples confirmed that TFA did not affect either retention times or separation and confirmed detected peaks as sucrose alkanoate regioisomers (Paper 1, Part II).

Regioisomer assignment was determined through structural analysis by ^{1}H- and ^{13}C-NMR spectroscopy (see table 2-2). Resonance assignments, identification of compounds and determination of alkanoate substitution positions were achieved from correlated spectroscopy (COSY) spectra and heteronuclear single quantum coherence (HSQC) spectra (Friebolin, 2005; Paper 1, Part II).

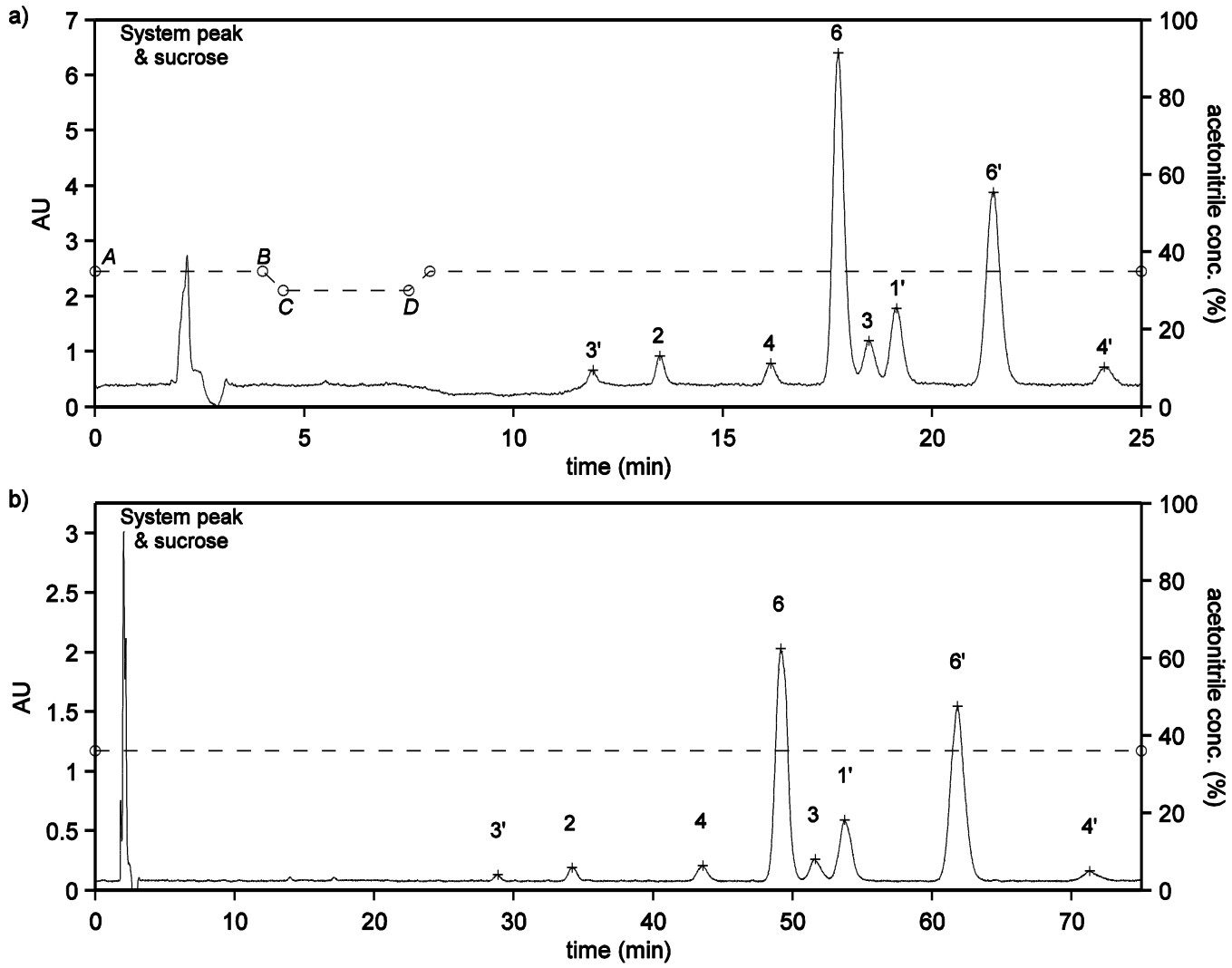


Figure 2-1: Chromatograms from RP-HPLC analyses of sucrose alkanoates, with profile of eluent acetonitrile concentration (--○). Peaks are annotated with regioisomer substitution position.
a) Sucrose caprate analysed with step-down elution profile described by the variables *A*, *B*, *C* and *D*: *A* = 35 %, *B* = 4 min, *C* = 30 %, *D* = 3 min.
b) Sucrose laurate analysed with isocratic elution at 36 % acetonitrile concentration.

Table 2-2: Chemical shifts of sucrose and sucrose caprate regioisomers in methanol-d_4 at 25 °C. Shifts are relative to internal tetramethylsilane. The chemical shifts of the C2' atoms were not recorded.

Atom	Sucrose	2	3	4	6	1′	3′	4′	6′
H1	5.39	5.53	5.44	5.44	5.38	5.40	5.39	5.42	5.35
C1	93.6	90.9	93.6	93.3	93.2	94.1	93.4	93.9	93.3
H2	3.42	4.60	3.58	3.52	3.45	3.42	3.42	3.45	3.45
C2	73.3	74.4	71.4	73.2	72.9	73.0	73.2	73.2	73.0
H3	3.70	3.88	5.21	3.88	3.74	3.68	3.61	3.73	3.74
C3	74.7	72.1	76.7	72.4	74.3	74.5	75.0	74.6	74.5
H4	3.35	3.42	3.52	4.81	3.52	3.37	3.39	3.38	3.36
C4	71.4	71.6	69.4	72.3	71.4	71.3	71.3	71.3	71.3
H5	3.82	3.89	3.91	4.02	4.03	3.83	3.81	3.88	3.84
C5	74.4	74.2	74.2	72.2	71.7	74.4	74.6	74.4	74.1
H6	3.80 3.71	3.82 3.71	3.81 3.73	3.60 3.49	4.39 4.18	3.80 3.71	3.80 3.74	3.84 3.74	3.84 3.73
C6	62.2	62.2	61.8	62.0	64.6	62.1	62.3	62.2	62.4
H1′	3.64 3.60	3.51 3.38	3.66 3.61	3.65 3.61	3.64 3.59	4.37 4.13	3.62 3.54	3.68 3.61	3.64
C1′	64.1	63.4	64.0	64.1	63.9	63.8	65.3	63.6	63.7
H3′	4.09	4.17	4.11	4.10	4.12	4.06	5.34	4.35	4.11
C3′	79.4	77.5	79.2	79.2	79.0	78.7	79.7	77.4	78.9
H4′	4.02	4.01	4.03	4.01	3.99	4.04	4.27	5.23	4.03
C4′	75.8	75.3	75.8	75.7	75.7	74.9	73.8	78.3	76.7
H5′	3.76	3.74	3.78	3.76	3.80	3.73	3.88	3.91	3.94
C5′	83.8	83.8	83.8	83.8	83.5	83.7	84.3	82.4	80.4
H6′	3.75	3.76	3.77	3.76	3.80 3.75	3.76	3.78	3.78 3.73	4.39 4.32
C6′	63.3	63.5	63.4	63.5	63.8	63.2	63.1	63.8	66.9

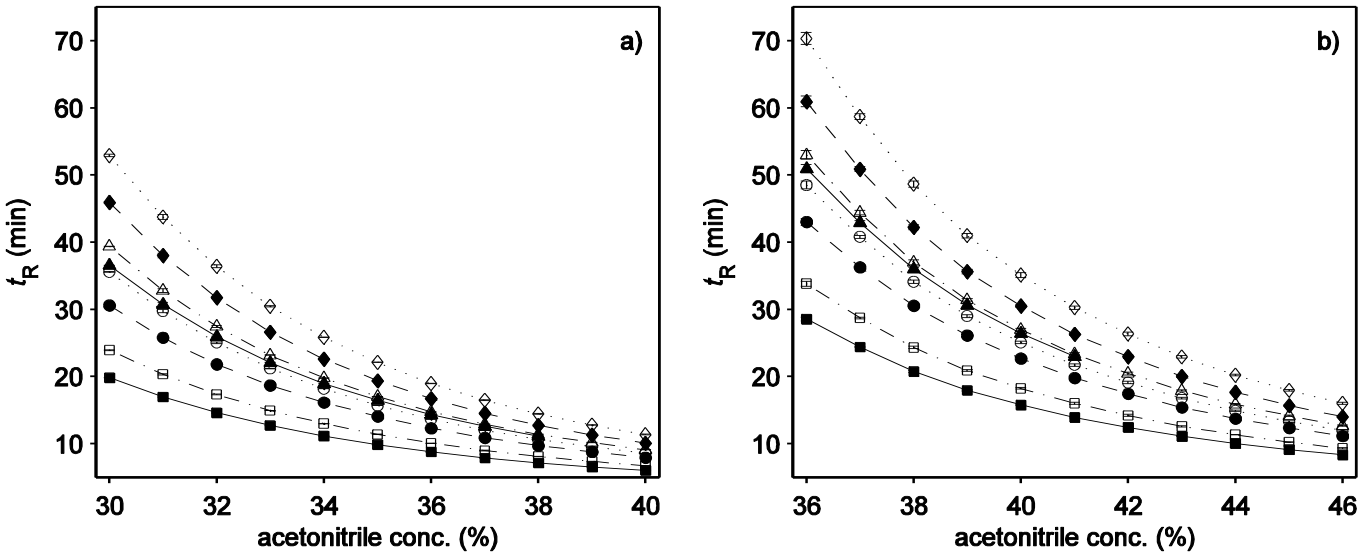


Figure 2-2: Regioisomer retention times as function of eluent acetonitrile concentrations for isocratic elutions of sucrose caprate (a) and sucrose laurate (b). The regioisomers are arrayed from bottom to top according to elution order: –■ 3', -·□ 2, --● 4, ··○ 6, –▲ 3, -·△ 1', --◆ 6', ··◇ 4'. Error bars indicate standard deviations.

2.4 Dependence of retention time on eluent acetonitrile concentration

The effect of eluent acetonitrile concentration on regioisomer retention times was investigated as an isolated factor over the ranges 30-40 % and 36-46 % acetonitrile for sucrose caprate and sucrose laurate, respectively. Increasing the eluent acetonitrile concentration decreased the retention times for all the sucrose alkanoate regioisomers over these ranges (see fig. 2-2). Also, at specific acetonitrile concentrations, increased retention times were observed for sucrose laurate regioisomers compared to the corresponding regioisomers of sucrose caprate, reflecting the difference in the fatty acid chain length between the sucrose alkanoates (Lisa et al., 2007).

Three regioisomers, 3, 6 and 1′, showed particularly similar elution properties for each of the sucrose alkanoates. The regioisomer 3-*O*-caproyl sucrose was resolved in the concentration range 30-38 %, while 3-*O*-lauroyl sucrose was resolved in the range 36-41 %, and above the indicated ranges these regiosiomers co-eluted with their respectively succeeding 1′-regioisomer. However, 3-*O*-lauroyl sucrose was resolved from the succeeding 1′-*O*-lauroyl sucrose at a lower acetonitrile concentration than that indicated by the retention time correspondence as compared to the concentration required for resolving the same pair of sucrose caprate regioisomers.

Similar, curved relationships between the eluent acetonitrile concentration and retention time were observed for all regioisomers of sucrose caprate and sucrose laurate, and were approximated well by cubic polynomials ($R^2 > 0.99$). Kendall rank correlation analysis confirmed the high interdependence of the curves (sucrose caprate regioisomers $\tau \geq 0.80$, sucrose laurate regioisomers $\tau \geq 0.89$, all regioisomers $\tau \geq 0.87$, see table 2-3). Slight variations in curvature between regioisomers were caused by experimental variations, and the lower τ-value for the 3-*O*-caproyl sucrose was attributed to the lower number of observations and the narrower concentration range over which this particular regioisomer was studied. As the curvature of the retention time dependence on eluent composition was the same for all regioisomers, the differences between their dependence functions were reflective of differences in physicochemical properties causing the specific elution sequence and separation of the sucrose alkanoate regioisomers in the HPLC-analyses.

Table 2-3: Kendall rank correlation coefficients (τ) for the second-order differences in retention times from isocratic elutions of sucrose caprate.
Values are the ratio of the difference between number of concordant and number of discordant rank order pairs to the total number of rank order pairs.

	2	**3**	**4**	**6**	**1′**	**3′**	**4′**	**6′**
2	*	4/5	1	1	1	1	1	1
3		*	4/5	4/5	4/5	4/5	4/5	4/5
4			*	1	1	1	1	1
6				*	1	1	1	1
1′					*	4/5	4/5	4/5
3′						*	1	1
4′							*	1
6′								*

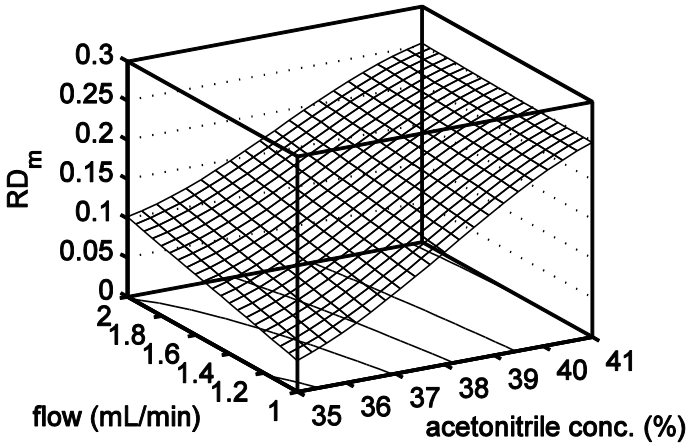


Figure 2-3: Response surface and contour plot for regression model of sucrose laurate general resolution deviation (RD_m) as function of eluent acetonitrile concentration and flow rate.
Regression function: $RD_m = 37.22 – 3.027I + 0.08091I^2 – 0.0007090I^3 + 0.3914F – 0.09414IF$;
$R^2 = 0.99$, $R^2_{adj.} = 0.99$; RMSE = 2.9 %. The range of RD_m was from 0 to 1, with 0 representing optimal resolution.

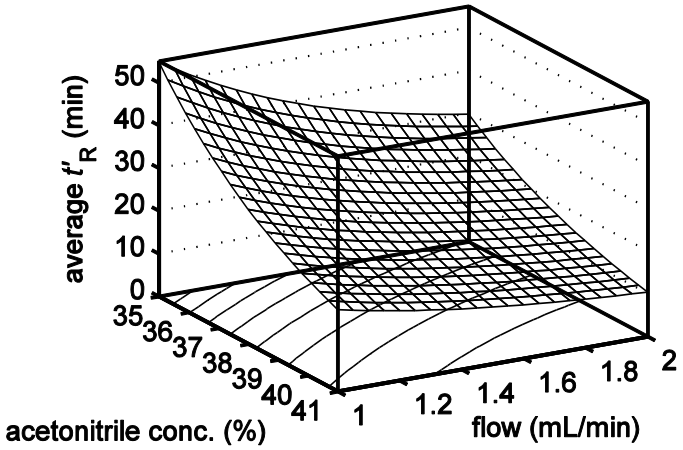


Figure 2-4: Response surface and contour plot for regression model of sucrose laurate average adjusted retention time ($\bar{t}'_R$) as function of eluent acetonitrile concentration and flow rate.
Regression function: $\log \bar{t}'_R = 15.84 – 0.4193I + 0.003205I^2 – 1.316F + 0.2350F^2$;
$R^2 = 0.98$, $R^2_{adj.} = 0.98$; RMSE = 0.38 %.

2.5 Isocratic elutions with varying flow rates

In analyses of sucrose laurate, flow rate effects under isocratic conditions were investigated using a 4-level factorial design with added centre point. The responses analysed were average adjusted retention time ($\bar{t}'_R$, min.) and general resolution deviation (RD_m).

Increasing the acetonitrile concentration and the flow rate resulted in increased RD_m (see fig. 2-3) and decreased $\bar{t}'_R$ (see fig. 2-4). This conformed to chromatographic theory, in that lower flow rates give better resolution, as caused by increased column efficiency (Tijssen, 1998). The experimental variations were estimated to RSD = 0.061 % for $\bar{t}'_R$, and RSD = 2.5 % for RD_m.

The statistical analysis and regression modelling showed that the effect of the acetonitrile concentration (variable *I*) on RD_m was approximately 2.5 times the effect of the flow rate (variable *F*), and while flow rate showed a linear effect, the effect of the acetonitrile

concentration followed a complex relationship, as characterised by statistically significant quadratic and cubic terms (see fig. 2-3). The interaction between the factors also showed a statistically significant effect on the resolution metric. The effect of the acetonitrile concentration on $\bar{t}'_R$ was approximately 1.6 times the effect of the flow rate, and both factors showed a curved relationship with the retention time response, however, the interaction between them was not statistically significant (see fig. 2-4).

Acetonitrile concentration and flow rate were both shown to be statistically significant for the resulting retention time and separation, however, changes in acetonitrile concentration (*I*) had larger effects than changes in flow rate (*F*) over the ranges studied. In addition, as the difference in effect between flow rate and acetonitrile concentration was larger for RD_m than for average adjusted retention time, the results indicated that any practical optimisation of separation and length of analysis time should be based on increased flow rates.

2.6 Step-down gradient elution profiles

The use of step-down gradient elution to improve regioisomer separation in the analysis of sucrose caprate was initially studied using a face-centred cubic (FCC) optimisation design with four variables, two concentration variables and two duration variables, as described above. The FCC design conformed to the limitations of the variable space and the curved relationship between eluent acetonitrile concentration and regioisomer retention times, as discussed above, made the use of an optimisation design necessary to model the design space adequately (Esbensen et al., 2002). Also, because the 3-regioisomer was resolved only within a limited variable space, the operational variable space was strictly delimited and corresponded to the region of interest, causing severe challenges in the application of circular experimental designs.

Table 2-4: Significance of selected variable effects, from PLS analysis of face-centred composite design, using cross-validation coefficient-variance significance testing.
Based on estimated *p*-values: +++ ($p < 0.005$), ++ ($p < 0.01$), + ($p \leq 0.05$), 0 ($p > 0.05$). Effects not included in the table were insignificant ($p > 0.05$). The significance levels for the exponential terms (A^x, C^x) were equal for all exponents in the range $2 \leq x \leq 4$.

	Responses							
Effect	**2**	**3**	**4**	**6**	**1′**	**3′**	**4′**	**6′**
A	+++	+++	+++	+++	+++	+++	+++	+++
B	0	0	0	0	0	0	0	0
C	+++	+++	+++	+++	+++	+++	+++	+++
D	0	0	0	0	0	0	0	0
AB	0	0	0	0	0	0	0	0
AC	+++	+++	+++	+++	+++	+++	+++	+++
BC	+	+	+	+	+	+	+	+
ABC	+++	+++	+++	+++	+++	+++	+++	+++
ACD	+	+	+	+	+	+	+	+
ABCD	+	+	+	+	+	+	0	0
A^x	+++	+++	+++	+++	+++	+++	+++	+++
C^x	+++	+++	+++	+++	+++	+++	+++	+++

The regioisomer retention times (t_R, min.) were analysed and modelled using partial least-squares (PLS) regression incorporating variable interaction effects between up to four design variables and exponential terms up to exponent 4 in the model (see Paper 1, Part II). The PLS-model was iteratively reduced using cross-validation coefficient-variance significance testing, until only significant effects remained. The experimental variation was estimated to RSD = 1.6 %.

The influence of the four elution profile variables on the resulting regioisomer retention times were shown to have varying significance (see table 2-4). The eluent acetonitrile concentrations were proved to have the highest significance, while a range of variable effects, including the individual terms for the duration variables, were determined as insignificant and could be disregarded. The concentration variables (*A* and *C*) were highly significant as isolated variables, as exponential terms, and as part of their interaction (*AC*). In addition, the interaction effect of the concentration variables and the duration of the initial concentration (*ABC*) was highly significant. The duration variables were only significant as part of interaction effects involving one or both of the concentration variables. The significance of the exponential terms of the concentration variables were interpreted as reflecting the nonlinear relationship between eluent acetonitrile concentration and retention time.

Table 2-5: Regression model evaluation.
Coefficients of determination (R^2 and $R^2_{adj.}$) and percent relative root-mean-square error (RMSE) for models of average adjusted retention time ($\bar{t}'_R$) and general resolution deviation (RD_m) for sucrose alkanoates.

	Sucrose caprate		Sucrose laurate	
Response	$\bar{t}'_R$	RD_m [a]	$\bar{t}'_R$	RD_m [b]
R^2	0.98	0.88	0.99	0.96
$R^2_{adj.}$	0.98	0.84	0.98	0.94
RMSE (%)	4.1	5.7	1.1	3.2

[a] Data transformed: log RD_m.
[b] Data transformed: $RD_m^{-1.15}$.

Building on and informed by the previously conducted experiments, the use of step-down gradient elution to improve regioisomer separation in the analysis of both sucrose caprate and sucrose laurate was investigated using circumscribed central composite (CCC) optimisation designs. The responses analysed were average adjusted retention time ($\bar{t}'_R$, min.) and general resolution deviation (RD_m). The data, untreated or transformed as necessary, were analysed by analysis of variance (ANOVA) and regression modelling on coded variables with the maximum number of exponential and interaction terms. The models were iteratively reduced, until only significant terms remained (see Paper 2, Part II).

In a number of analyses of sucrose caprate the 2- and 3′-regioisomers eluted during the step-down gradient, that is, satisfying the relation $t_R < (B + D + 1 + t_0)$. This introduced an additional factor of experimental variation for these regioisomers, resulting in RSD = 24 % for RD_m, however, the effect on $\bar{t}'_R$ was lower, resulting in RSD = 2.4 %. The experimental design for sucrose laurate avoided such elution during the step-down gradient by constraining the maximum summed time of the duration variables *B* and *D*, resulting in lower experimental

variation for both RD_m and $\bar{t}'_R$, which showed RSD = 2.2 % and RSD = 1.2 %, respectively. All regression models explained more than 88 % of the variation in the experiments, but most models explained more than 96 % of the variation observed (see table 2-5). The modelling errors were also of the same order of magnitude as, or lower than, the experimental variations.

The regression models for step-down gradient elution of the sucrose alkanoates showed that the eluent acetonitrile concentrations were the overall most significant variables for retention time and separation. The models for average adjusted retention time were very similar in the number of terms, but differed in the significance of four terms (see table 2-6). The regression model for $\bar{t}'_R$ of sucrose caprate included the additional cubic term for the concentration variable *C* at a low significance level. The linear term for variable *B* was not significant, while all other terms were highly significant. The corresponding model for sucrose laurate included only significant terms, of which all but two were highly significant, the exceptions being the quadratic term for the concentration variable *A* and the interaction between the successive concentration and its duration (*CD*), which both showed a low level of significance. Although there were no variables clearly dominating in significance in these models, there was a distinction in the effects between the concentration variables (*A* and *C*) and the duration variables (*B* and *D*), where the former were included as linear terms, exponential terms and part of interaction terms, while the latter variables only contributed as linear or interaction effects. The models were also characterised by simple interactions, in the sense that the only significant interactions were between each concentration variable and the connected duration variable.

The regression models for RD_m showed larger differences between the sucrose alkanoates, in the number of terms and their significance (see table 2-6). The regression model for RD_m of sucrose caprate consisted of 14 terms at varying levels of significance, but was characterised by the concentration variables (*A* and *C*) being the most significant. These variables (*A* and *C*) were highly significant as linear and quadratic terms, while the cubic term for *A* showed low significance. The linear terms for duration variables (*B* and *D*) were not significant, while the quadratic term for *D* showed low significance. Of the included two-variable interactions only the interaction between the initial concentration and its duration (*AB*) was significant at medium level, while the three-variable interaction *ACD* showed low significance. The complexity of this regression model was interpreted as arising from the issue of the experimental variation exhibited by the sucrose caprate RD_m (RSD = 24 %), as also reflected in the lower coefficients of determination for this model compared to the other regression models. The corresponding model for sucrose laurate consisted of 10 terms, also at varying significance levels, and characterised by the initial concentration and its duration (*A* and *B*) being the most significant variables. The linear terms for all variables were significant, with the initial concentration and its duration (*A* and *B*) at high level and the successive concentration and its duration (*C* and *D*) at low level. The quadratic terms for *A* and *B* were not significant, the cubic terms for these variables showed low significance, while their mutual interaction (*AB*) was highly significant.

Step-up gradient elutions were performed by modification of step-down profiles by exchanging the values of the variables *A* and *C* and changing the variables *B* and *D* to the respective mean values. The step-up gradients resulted in similar average adjusted retention times and general resolution deviations as the step-down profiles.

Table 2-6: Coefficients and significance of terms for regression models for average adjusted retention time ($\bar{t}'_R$) and general resolution deviation (RD_m) for sucrose alkanoates.
Significance levels: *ns* $p > 0.1$, * $p \leq 0.1$, ** $p < 0.01$, *** $p < 0.001$.

	Sucrose caprate		Sucrose laurate	
Response	$\bar{t}'_R$	RD_m[a]	$\bar{t}'_R$	RD_m[a]
Constant	-48.19	143.8	-78.22	738.3
A	4.078 (***)	-10.76 (***)	4.643 (***)	-52.33 (***)
B	6.370 (*ns*)	-1.328 (*ns*)	10.76 (***)	2.726 (***)
C	-0.8306 (***)	-0.7061 (***)	1.351 (***)	-3.310×10^{-2} (*)
D	2.407 (***)	-5.369 (*ns*)	2.411 (***)	0.1121 (*)
A^2	-6.148×10^{-2} (***)	0.2772 (***)	-5.927×10^{-2} (*)	1.256 (*ns*)
A^3		-2.432×10^{-3} (*)		-1.007×10^{-2} (*)
B^2				0.2714 (*ns*)
B^3				2.603×10^{-2} (*)
C^2	4.418×10^{-2} (***)	5.643×10^{-3} (***)	-2.338×10^{-2} (***)	
C^3	-7.069×10^{-4} (*)			
D^2		4.794×10^{-2} (*)		
AB	-0.1849 (***)	3.773×10^{-2} (**)	-0.2784 (*)	-9.280×10^{-2} (***)
AC		1.113×10^{-2} (*ns*)		
AD		0.1437 (*ns*)		
CD	-6.590×10^{-2} (***)	0.1721 (*ns*)	-5.527×10^{-2} (*)	
ACD		-4.900×10^{-3} (*)		

[a] Data transformed. See table 2-5 for details

2.7 Evaluation of elution efficiency

The efficiency of the investigated elution strategies was evaluated in terms of RD_m and analysis time (see fig. 2-5), where the best results were offered by the elution strategy minimising both of these values for each of the two sucrose alkanoates (Kromidas, 2006). In the analysis of sucrose caprate the most efficient elutions were achieved with step-down gradient elution, while for sucrose laurate the most efficient elutions were achieved using isocratic elution with increased flow.

The sucrose caprate analyses (see fig. 2-5a) showed the best results for the isocratic elutions in the concentration range 33-35 % (analysis time 32, 27 and 24 minutes, respectively). The lowest RD_m was achieved at 33 %, with $RD_m = 0.060$, however, elution at 34 % reduced the analysis time by 16 % with only a 15 % increase in RD_m, to 0.069. Step-down gradient elution offered improved separation at similar analysis time compared to isocratic elution, as the majority of the step-down gradient analyses resulted in analysis times between 25 and 31 minutes, of which twelve elutions showed $RD_m < 0.07$. The lowest values for RD_m were obtained with the acetonitrile concentration variables at 32.5 % and 25 % (*A* and *C*, respectively) resulting in RD_m in the range 0.052-0.037 and analysis time in the range 29-31 min, corresponding to 3-9 %

reductions in analysis time and RD_m decreased by 13-38 % relative to isocratic elution at 33 %. These highly similar results for several value combinations of the duration variables were interpreted as a manifestation of the lower statistical significance of these variables, as shown by the regression analysis.

The efficiency evaluation for analyses of sucrose laurate (see fig. 2-5b) showed that the step-down gradient elution strategy that offered improved efficiency for sucrose caprate saw a diminishing return when transferred to sucrose laurate, even to the point that isocratic conditions offered better efficiency. Step-down and step-up gradient elutions exhibited generally poor efficiency, with values of RD_m above 0.14 and analysis times above 43 minutes, while among the isocratic elutions, which exhibited a strong negative linear correlation between RD_m and analysis time ($r = -0.99$), the acetonitrile concentrations 36 % ($RD_m = 0.052$, 72 min), 37 % ($RD_m = 0.10$, 60 min) and 38 % ($RD_m = 0.14$, 50 min) offered the most efficient results. For isocratic elutions with varying flow the best efficiency was obtained with flow rates in the range 1.33-2.0 mL/min and concentrations of 35 or 37 % acetonitrile, which resulted in analysis time from 33 to 67 minutes and $RD_m < 0.15$. For these elutions there was also observed a strong negative linear correlation ($r = -0.99$) between RD_m and analysis time. Compared with the results for elution at 38 % acetonitrile, the improvements over isocratic conditions with standard flow (1.0 mL/min) were reduction of analysis time by 34 % with equal RD_m for elution at 37 % with flow 2.0 mL/min, while elution at 35 % with flow 2.0 mL/min reduced RD_m by 29 % and analysis time by 6 %.

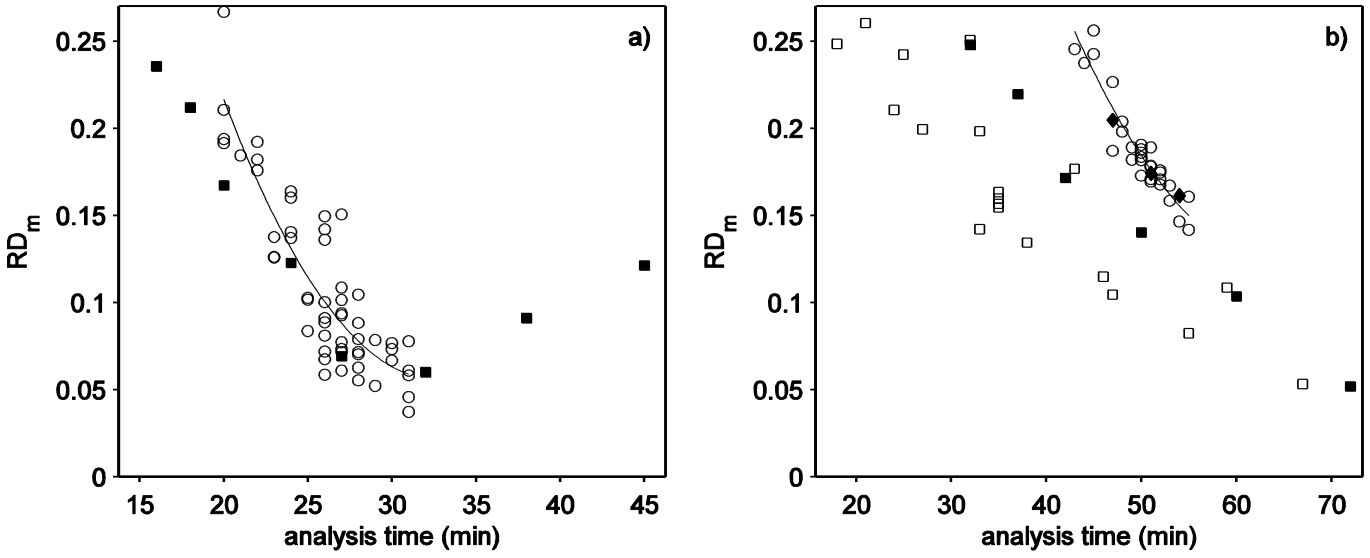


Figure 2-5: General resolution deviation (RD_m) and analysis time from elution strategies for sucrose caprate (a) and sucrose laurate (b). Elutions: Isocratic (■); different flow rates under isocratic conditions (□); step-down gradient profiles (○); and step-up gradient profiles (◆). Lines show the quadratic polynomial fits of the step-down gradient profile elutions, with $R^2 = 0.79$ (a) and $R^2 = 0.89$ (b).

3 SYNTHESIS OF SUCROSE LAURATE IN *N,N*-DIMETHYLFORMAMIDE

The development of methods for quantitative analysis of sucrose alkanoates using RP-HPLC (described in ch. 2) facilitated the study of enzyme-catalysed ester syntheses with increased detail and accuracy and, in particular, offered increased resolution and sensitivity in the analysis of regioisomeric distribution. The irreversible acylation of sucrose with the enol ester vinyl laurate as acyl donor catalysed by serine proteases and a metalloprotease in the hydrophilic, aprotic solvent DMF was chosen for detailed study (see sections 1.2.1 and 1.2.2).

Catalytic activity and appearance and distribution of monoester regioisomers were investigated in enzyme-catalysed reactions and control reactions with non-catalytic protein and without protein. Reactions were performed with the enzyme formulations thermolysin, ALP-901, subtilisin A and alcalase 3.0T, with the non-catalytic protein casein and without protein, as detailed in Paper 3 (Part II). Standard curves were constructed for quantitative RP-HPLC analysis of the reaction mixture, and sucrose laurate regioisomers were analysed by mass spectrometry (MS).

3.1 Quantitative RP-HPLC analysis

Sample series of sucrose and 6-*O*-lauroyl sucrose were analysed to construct standard curves for quantitative RP-HPLC with CAD. Obtained peak areas were plotted against injected mass on logarithmically scaled axes (see fig. 3-1 and fig. 3-2) to determine the linear mass-response relation for each analyte, as described in section 1.3.2.

Linear regression resulted in slopes and intercepts according to eq. 1-4, which were inserted in eq. 1-5 to obtain equations for determining injected mass of analyte (*M*, ng) from peak area (*A*, pA*s) for sucrose

$$M \pm 2.1\ \% = (A/a)^{1/b} = (A/1.38)^{1.39} = \frac{A^{1.39}}{1.565} \tag{3-1}$$

and sucrose laurate

$$M \pm 3.5\ \% = (A/a)^{1/b} = (A/0.274)^{1.09} = 4.20A^{1.09} \tag{3-2}$$

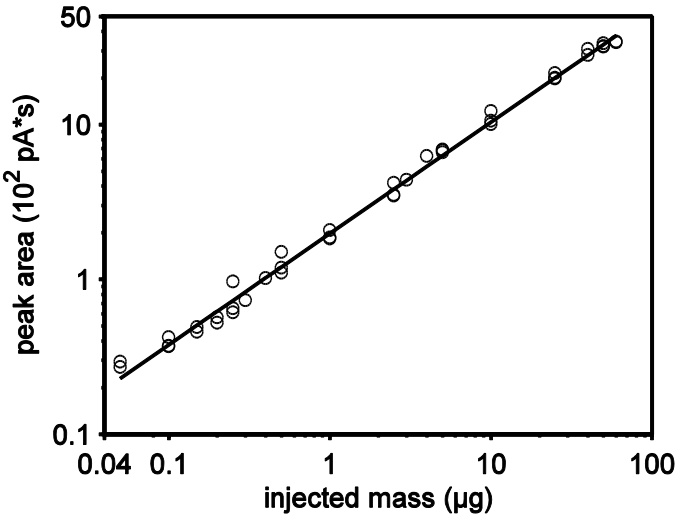


Figure 3-1. Standard curve for sucrose in RP-HPLC analysis with CAD.
Log-log plot of results from analysis of sucrose sample series corresponding to the working ranges 0.05-60 µg and 30-3750 pA*s, with linear regression line.
Coefficients of determination for linear regression: R^2 = 0.995, $R^2_{adj.}$ = 0.995.

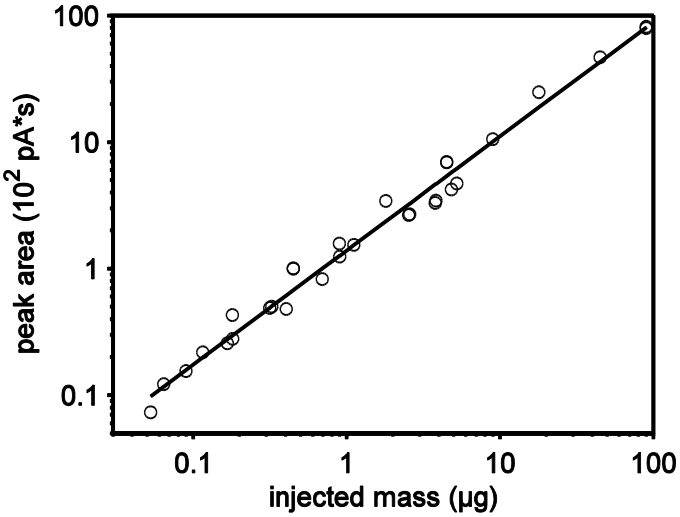


Figure 3-2. Standard curve for sucrose laurate in RP-HPLC analysis with CAD.
Log-log plot of results from analysis of 6-*O*-lauroyl sucrose sample series corresponding to the working ranges 0.06-90 µg and 10-8000 pA*s, with linear regression line.
Coefficients of determination for linear regression: R^2 = 0.983, $R^2_{adj.}$ = 0.982.

3.2 Mass spectrometry of sucrose laurate regioisomers

Isolated samples of 6- and 6′-*O*-lauroyl sucrose were analysed using mass spectrometry (MS). Matrix-assisted laser desorption/ionisation (MALDI) time-of-flight MS resulted in signals characteristic for the molecular ion adducts with sodium and potassium ($[M+Na]^+$ and $[M+K]^+$, *m*/*z* 547 and 563, respectively; see fig. 3-3), which corresponded to the observations reported by Wang et al. (2007a) for glucose, fructose and sucrose. Additional signals with relative intensity above 10 % were observed for fragment ions of both regioisomers. Taylor et al. (2005) reported that sucrose, glucose and fructose exhibited extensive fragmentation pathways with scission of the glycosidic bond (for sucrose) and neutral losses of the molecular units CH_2O and H_2O in MS with soft ionisations techniques. As most of the signals observed in the MALDI spectra precluded the scission of the ester bond, it was assumed that sucrose laurate predominantly exhibited corresponding fragmentation pathways of the sucrose group.

The signals common to both regioisomers located at *m*/*z* 345 and 273 respectively corresponded to scission of the glycosidic bond to form either of the glycosyl ester ions $[M–Glc+H]^+$ or $[M–Fru+H]^+$, and the subsequent loss of glyceraldehyde or dihydroxyacetone ($C_3H_6O_3$) to form $[M–C_9H_{18}O_9+H]^+$. Two of the unique signals observed for 6-*O*-lauroyl sucrose resulted from fragmentation without scission of the glycosidic bond (see fig. 3-3a). The signal located at *m*/*z* 439 corresponded to the ion adduct with sodium after loss of 3 CH_2O molecules and 1 H_2O molecule, most likely all from the fructosyl group, to form $[M–C_3H_8O_4+Na]^+$. The signal at *m*/*z* 409 corresponded to an additional loss of CH_2O, resulting in the ion adduct $[M–C_4H_{10}O_5+Na]^+$, with the most likely pathway being fragmentation of H_2O from the glucosyl, and $C_4H_8O_4$ from the fructosyl, resulting in an acetaldehyde group. The signal at *m*/*z* 299 could not be accounted for with the above fragmentation pathways while preserving the ester bond, however, with scission of the ester, the subsequent loss of 1 CH_2O molecule and 2 H_2O molecules from either of the glycosyl groups would form the ion adduct with sodium $[M–C_{13}H_{28}O_4+Na]^+$ at the observed mass.

Electrospray-ionisation (ESI) tandem mass spectrometry (MS/MS) on the precursor ion $[M+Na]^+$ was performed to investigate whether the regioisomers of sucrose laurate could be differentiated using MS analysis. The resulting product ion mass spectra exhibited two high-intensity signals for each sample (see fig. 3-4). The signals for 6-*O*-lauroyl sucrose were located at *m*/*z* 547, corresponding to the precursor ion $[M+Na]^+$, and at *m*/*z* 385, corresponding to the 6-*O*-lauroyl glucose ion adduct with sodium $[M−C_6H_{10}O_5+Na]^+$, resulting from scission of the glycosidic bond (see fig. 3-4a). The signals for 6′-*O*-lauroyl sucrose were located at *m*/*z* 385 and 367, respectively corresponding to scission of the glycosidic bond on either side of the linking oxygen, forming the 6-*O*-lauroyl fructose ion adduct with sodium $[M−C_6H_{10}O_5+Na]^+$ or the fructosyl ester ion adduct with sodium $[M−Glc+Na]^+$ (see fig. 3-4b).

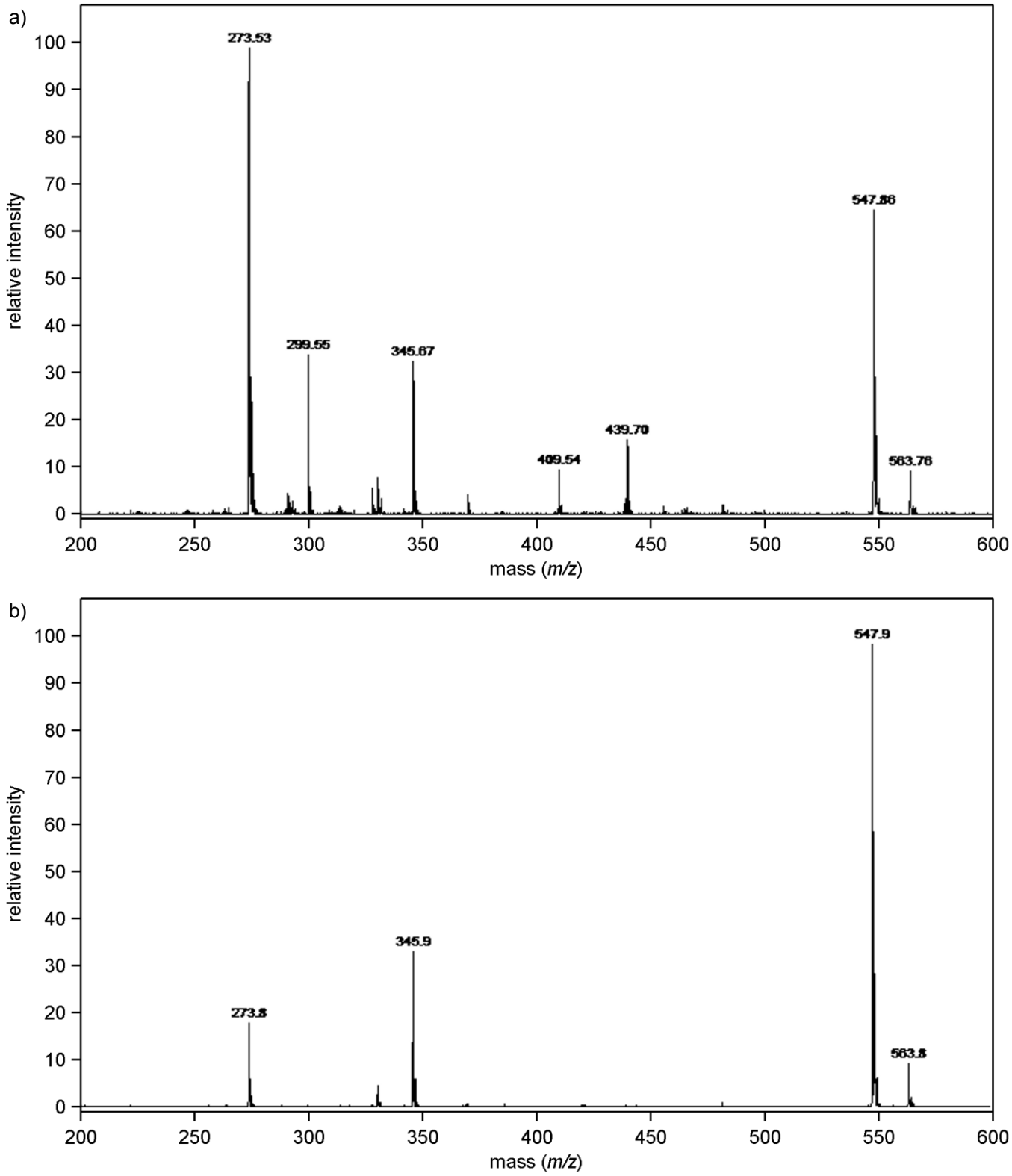


Figure 3-3. MALDI mass spectra of 6-*O*-lauroyl sucrose (a) and 6′-*O*-lauroyl sucrose (b). MS analysis was performed using 2,5-dihydroxybenzoic acid as matrix (see Paper 3, Part II).

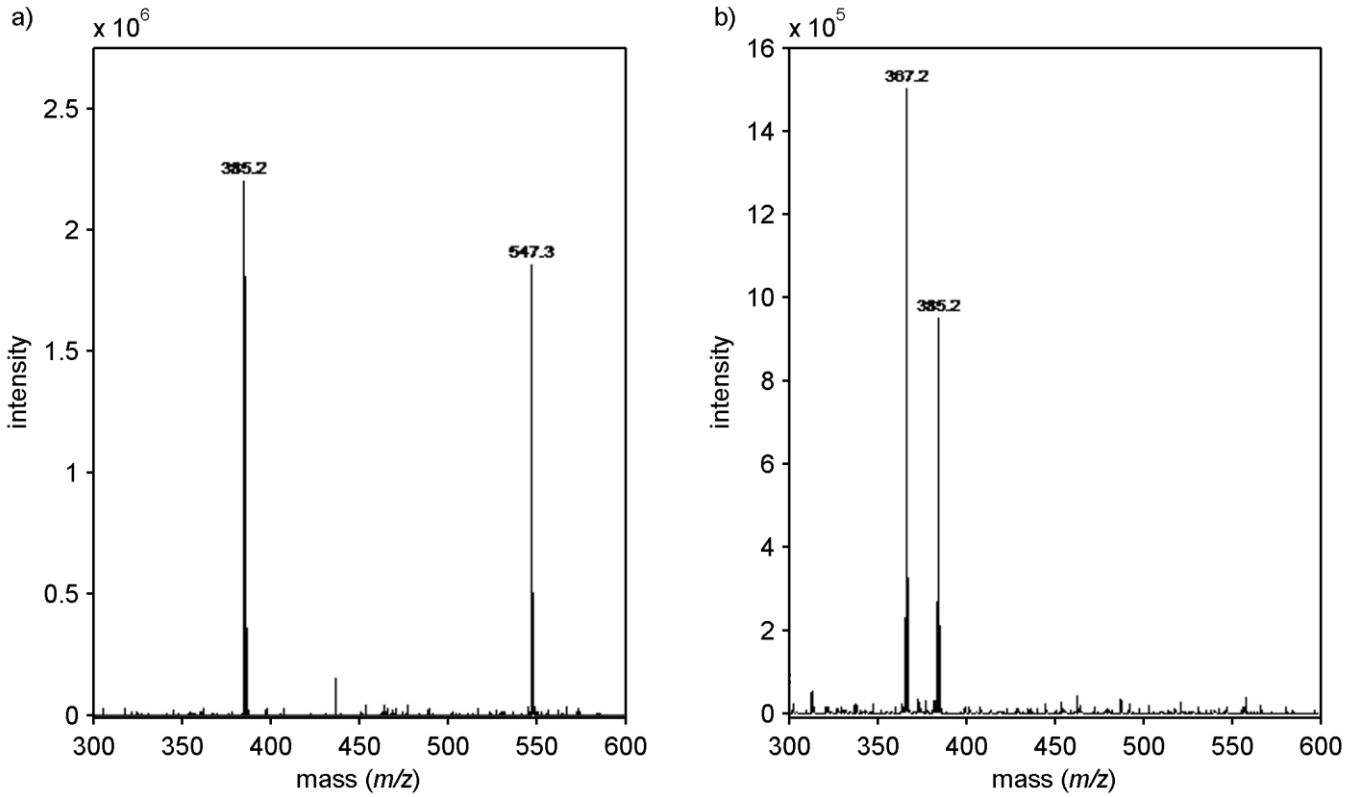


Figure 3-4. Product ion mass spectra from ESI-MS/MS on precursor ion $[M+Na]^+$ of 6-*O*-lauroyl sucrose (a) and 6′-*O*-lauroyl sucrose (b).
Samples (5.5 g/L in MeOH-ACN-H_2O [10:35:15, *v*/*v*] w/0.025 % [*v*/*v*] formic acid) were analysed by nanoelectrospray-ionisation high-capacity-ion-trap tandem mass spectrometry with collision-induced dissociation on the precursor ion.

These results indicated a difference in stability of the molecular ion between the regioisomers, and a difference in the preferential scission of the glycosidic bond, with the 6-regioisomer losing the fructosyl radical, while the 6′-regioisomer exhibited a preference for loss of the glucose molecule over the glucosyl radical at about a 3:2 ratio. However, as the regioisomers analysed were acylated at corresponding primary hydroxy positions in either saccharide residue, it cannot be determined whether MS/MS product ion mass spectra can be used for identification of the specific regioisomers, or if the extent of difference discernable by MS is whether the ester substitution is on the glucose or the fructose residue. Further study of sucrose laurate regioisomers acylated at secondary hydroxy positions and on the same saccharide residue analysed by mass spectrometry is necessary to draw firm conclusions.

3.3 Synthesis of sucrose laurate regioisomers

All investigated reactions, including the control reactions with casein and without protein, resulted in the synthesis of sucrose laurate. The concentration of sucrose was observed to decrease progressively during the course of all the reactions, and the total conversion of sucrose after 48 hours ranged from 19 to 96 % (see table 3-1). The highest degree of conversion was observed in the reaction with thermolysin, while the reactions with no protein and ALP-901 resulted in degrees of conversion above 50 %. The sucrose laurate yield after 48 hours ranged from 12 to 53 %, with the reaction catalysed by ALP-901 resulting in the highest yield. In all reactions, 2-*O*-lauroyl sucrose was the monoester synthesised in highest abundance and its concentration increased over the first 6 hours of reaction (see fig. 3-5).

Formation of di- and oligoesters at the expense of monoesters was indicated by the comparison of yield of sucrose laurate with degree of substrate conversion in the reactions with thermolysin, ALP-901 and without protein. The presence of oligoesters was observed after 48 hours in these reactions, and the total detected amounts of esters corresponded to the degree of sucrose conversion.

Table 3-1. Sucrose conversion and sucrose laurate yields after 48 hours.
Conversion expressed as ratio of concentration decrease to initial concentration, and yield expressed as ratio of detected concentration to theoretical maximum.

Reaction	Sucrose conversion (%)	Sucrose laurate yield (%)
Thermolysin	96 ± 3	29 ± 2
ALP-901	66 ± 5	53 ± 4
Alcalase	34 ± 5	44 ± 5
Subtilisin	19 ± 8	12 ± 3
Casein	21 ± 9	34 ± 9
No protein	82 ± 2	46 ± 10

3.3.1 Catalytic activity

The observed initial rates of formation for 2-*O*-lauroyl sucrose ranged from 0.765 to 108 µM/min (see table 3-2) and the conditions exhibiting highest initial rates were thermolysin, ALP-901 and no protein. The rates with thermolysin and ALP-901 were, respectively, about 33 and 7 times the rate for the reaction without protein.

The reaction without protein appeared to be catalysed by the molecular sieves employed in the investigation. This reaction exhibited an initial rate about 2-4 times that observed with alcalase, casein and subtilisin, indicating the presence of catalytic functionality other than enzyme under these reaction conditions. Diatomaceous earth (Celite) was reported to exhibit catalytic activity in the esterification of sucrose with vinyl stearate in DMF/DMSO (1:1, *v*/*v*) by Ritthitham et al. (2009b), and the molecular sieves employed in the investigated reactions were of sodium aluminosilicate, a material similar to Celite. In medium without Celite the initial rate of formation for 2-*O*-stearoyl sucrose was 0.17 µM/min, while it increased to 1.7 µM/min with 10 g/L Celite in the medium and to 11.7 µM/min at 100 g/L. The initial rate observed in the reaction without protein (3.24 µM/min) and the concentration of molecular sieves (40 g/L) corresponded well with these observations.

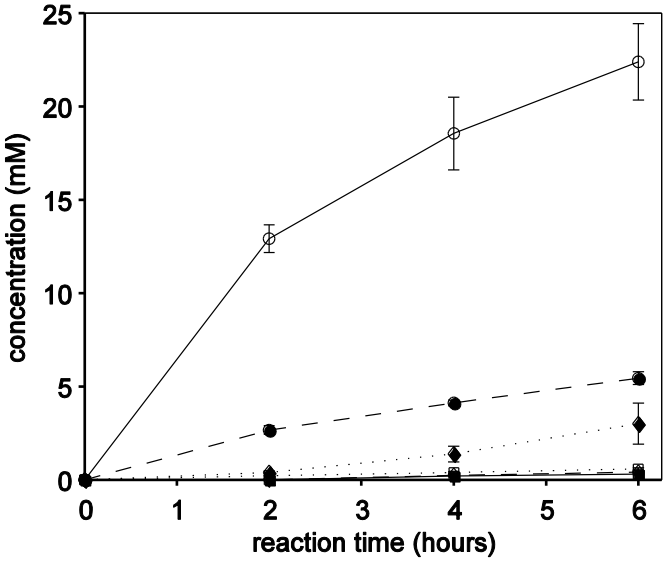


Figure 3-5. Progress curves for 2-*O*-lauroyl sucrose.
Reaction conditions: –○ thermolysin, --● ALP-901, ···□ alcalase, –■ subtilisin, --◇ casein, ···◆ no protein. Error bars indicate standard deviation determined from triplicate reactions, except for thermolysin (duplicates).

Addition of the non-catalytic protein casein to the reaction seemed to lower the catalytic effect of the molecular sieves, possibly due to interaction with the protein, as indicated by the decreased initial rate of formation compared to the reaction without protein. The initial rates with the subtilisin formulations were the same order of magnitude as that with casein, indicating similarly decreased catalytic activity in these reactions. Substrate inhibition, particularly with vinyl laurate, could decrease the activity of the enzymes, and the tautomerisation by-product in acyl transfer with vinyl esters, acetaldehyde, can act as an alkylating agent through formation of Schiff's bases with lysine residues in the protein, which destabilises some enzymes (Weber et al., 1995). However, subtilisin has been reported to catalyse acyl-transfer reactions with vinyl esters at higher acyl donor concentrations than employed in the investigated reactions (Pedersen et al., 2003) and the presence of molecular sieves has been indicated as beneficial for enzyme stability in enol ester acylations through trapping of acetaldehyde (Faber, 2004). As subtilisin exhibits high stability in DMF (see section 1.2.2), known causes could not account for decreased catalytic activity, and the remaining conclusion appeared to be that catalytic activity specifically ascribable to the subtilisin formulations was not observed in either reaction.

Table 3-2. Initial rates of formation for 2-*O*-lauroyl sucrose.
Initial rates (μM/min) were calculated from 2-hour samples, except for the casein reaction where it was calculated from the 4-hour sample.

Enzyme formulation				Controls	
Thermolysin	**ALP-901**	**Alcalase**	**Subtilisin**	**Casein**	**No protein**
108 ± 6	22.2 ± 1.9	1.78 ± 0.17	0.765 ± 0.234	0.981 ± 0.020	3.24 ± 0.63

3.3.2 Distribution and appearance of sucrose laurate regioisomers

Different concentrations of sucrose laurate regioisomers were observed in the investigated reactions after 48 hours, with 4 to 6 regioisomers detected in quantifiable amounts (see fig. 3-6). The regioisomeric distribution after 48 hours was similar in the reactions with alcalase, subtilisin and casein, while it differed somewhat in the reactions with ALP-901 and without protein, whose distributions were similar to each other. Common for all these reactions were, in descending order of abundance, the presence of the 2-, 3′- and 1′-regioisomers (see table 3-3). The reaction with thermolysin exhibited a unique regioisomeric distribution with lower proportion of the 3′-regioisomer and the 4-regioisomer not detected. The detected appearance of the sucrose laurate regioisomers largely corresponded to the apparent rates of formation, based on regioisomeric distribution and concentration after 48 hours.

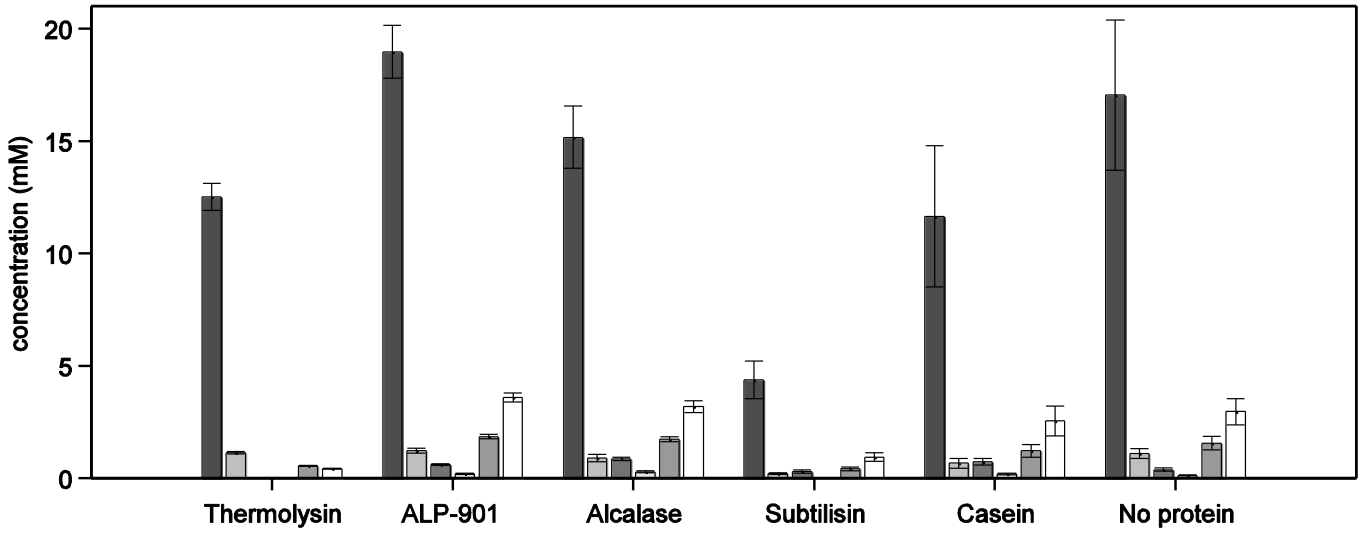


Figure 3-6. Catalyst effect on sucrose laurate regioisomer concentration in reaction mixture after 48 hours. Regioisomers of sucrose laurate: ■ 2, ■ 3, ■ 4, ■ 6, ■ 1′, □ 3′. The 4′- and 6′-regioisomers were not detected except in trace amounts in the reaction without protein. Error bars indicate standard deviations, determined from triplicate reactions, except for thermolysin (duplicates).

Table 3-3. Catalyst effect on appearance and distribution of sucrose laurate regioisomers. Regioisomer identity is indicated under the reaction time of first detection for each reaction condition.

Reaction	Reaction time					Regioisomeric distribution after 48 hours (2:3:4:6:1′:3′, *c*/*c*)[a]
	2 h	4 h	6 h	24 h	48 h	
Thermolysin	2, 3, 4, 1′, 3′					86:8:-:-:4:3
ALP-901	2, 4, 1′, 3′	3, 6				72:5:2:1:7:14
Alcalase	2	1′, 3′		3, 4, 6		69:4:4:1:8:14
Subtilisin	2		3’	4, 1′	3	71:3:5:-:7:15
Casein		2, 3′		3, 4, 1′	6	69:4:4:1:7:15
No protein	2, 3′	3, 4, 6, 1′		[b]	4′, 6′	74:5:2:1:7:13

[a] The regioisomers 4′ and 6′ were only detected in trace amounts in the reaction without protein. Sums of ratios can differ from 100 due to rounding.

[b] The reaction without protein was not sampled after 24 hours.

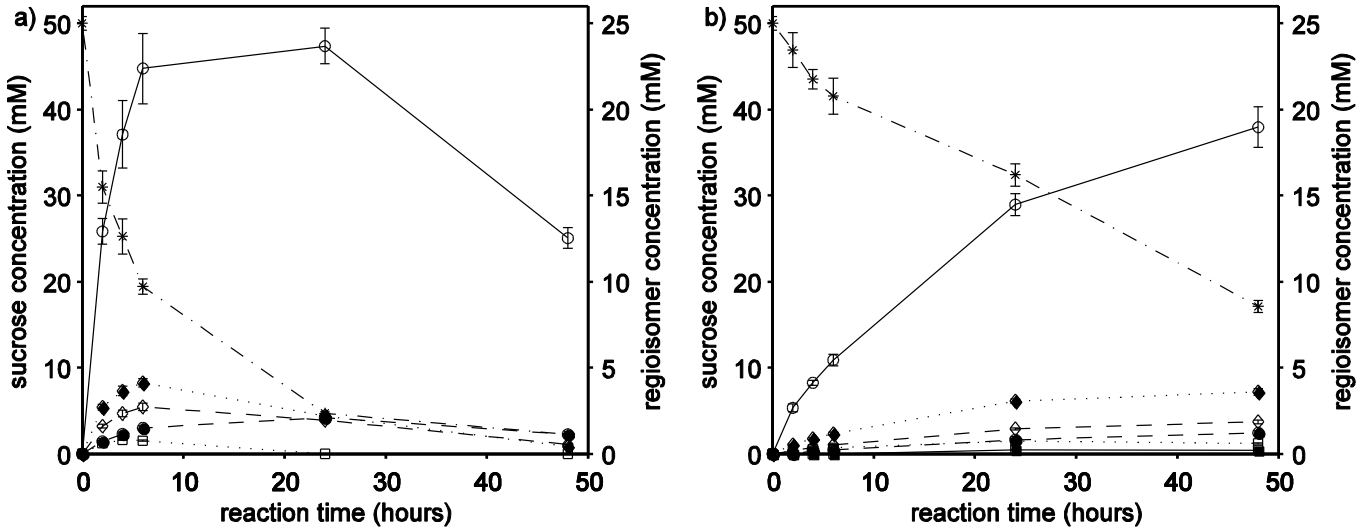


Figure 3-7. Progress curves for the reactions catalysed by thermolysin (a) and ALP-901 (b).
Sucrose: -·✳; sucrose laurate regioisomers: –○ 2, --● 3, ···□ 4, –■ 6, --◇ 1′, ···◆3′.
Error bars indicate standard deviations determined from duplicates for thermolysin and triplicates for ALP-901.

Comparison of reaction progress over time for the reactions with thermolysin and ALP-901 revealed that the reaction progress over 48 hours with ALP-901 resembled that over 4-6 hours with thermolysin (see fig. 3-7). In the thermolysin-catalysed reaction the highest concentration of 2-*O*-lauroyl sucrose (23.7 mM) was observed after 24 hours (see fig. 3-7a). The 4-regioisomer concentration was highest after 4 hours and could no longer be detected from 24 hours. The concentrations of regioisomers 1′ and 3′ was highest after 6 hours, while the concentration of the 3-regioisomer was highest after 24 hours. The overall highest yield of sucrose laurate (63 ± 5 %) was detected in this reaction after 6 hours, where the regioisomeric distribution was 71:5:2:-:9:13 (2:3:4:6:1′:3′, *c/c*) and the degree of sucrose conversion indicated only monoester product was present. In the reaction catalysed by ALP-901 the 2-*O*-lauroyl sucrose concentration increased progressively during the reaction to 19.0 mM after 48 hours (see fig. 3-7b). The concentration of regioisomers 4 and 6 was highest after 24 hours, while the other regioisomers detected were observed in highest concentration after 48 hours of reaction.

The reaction progress over 6 hours for the control reaction without protein was also considered (see fig. 3-8). The 2-regioisomer was the most abundant at 3.00 mM after 6 hours, and the regioisomeric distribution at this time was 69:3:5:1:7:16 (2:3:4:6:1′:3′, *c*/*c*). After 48 hours, all regioisomers were detected at the highest concentrations observed in this reaction, with 17.0 mM of 2-*O*-lauroyl sucrose detected.

The differences in regioisomeric distribution after 48 hours, such as the distinct distribution in the reaction with thermolysin, appeared partly to stem from differences in the overall reaction rates. The regioisomeric distribution after 6 hours in the reaction without protein was similar to the distributions in the reactions with alcalase and casein after 48 hours, and the distribution after 6 hours in the reaction with thermolysin was similar to the regioisomeric distribution in the reaction with ALP-901 after 48 hours (see table 3-3). In the two latter reactions, similar degrees of sucrose conversion were observed at the respective reactions times indicated, and the concentration of some of the same regioisomers were observed to decline in concentration from after 4 hours with thermolysin and after 24 hours with ALP-901.

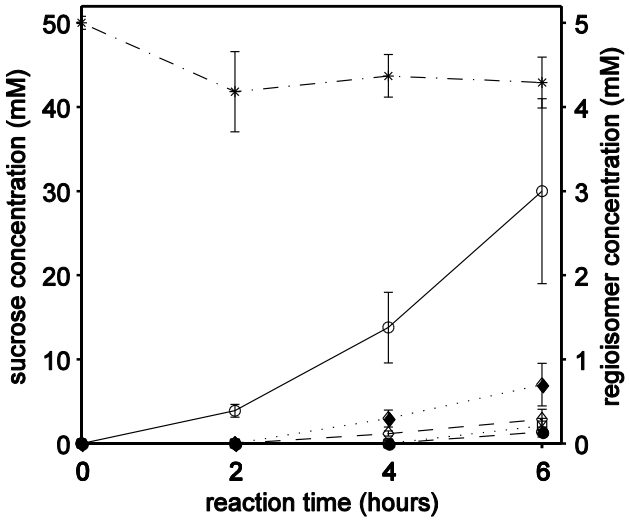


Figure 3-8. Progress curves for the reaction without protein.
Sucrose: -·✳; sucrose laurate regioisomers: –○ 2, --● 3, ···□ 4, –■ 6, --◇ 1′, ···◆3′.
Error bars indicate standard deviations determined from triplicate reactions.

3.3.3 Regioselectivity in the sucrose laurate syntheses

Thermolysin and ALP-901 exhibited regioselectivity predominantly towards formation of the 2-regioisomer of sucrose laurate in the investigated acylation reactions. This corresponded to previous reports on thermolysin-catalysed acyl transfer reactions in hydrophilic, aprotic solvents (Perez-Victoria and Morales, 2006), and reports on catalysis by another alkaline protease, AL-89, in vinyl fatty acid acyl transfers to sucrose (see section 1.2.2). Previous reports also suggested that employing ratios of acyl donor to acyl acceptor lower than 8:1 could increase the selectivity for monoester product (Ferrer et al., 1999; Pedersen et al., 2003). Previously reported subtilisin-catalysed reactions with vinyl esters and activated esters in DMF and other solvents have exhibited selectivity towards acylation at the 1′-position of sucrose (see section 1.2.2), however, as outlined above, the catalytic activity of the subtilisin formulations appeared negligible under the investigated reaction conditions.

In the reactions determined to exhibit catalytic activity from the molecular sieves, the regioisomeric distribution appeared to be determined by solvent influence. As previously discussed, in hydrophilic, aprotic solution, sucrose is conformation-stabilised by intramolecular hydrogen bonds between 2-OH and either of 1′-OH and 3′-OH, resulting in the hydroxy group reactivity order 2-OH > 1′-OH > 3′-OH (see sections 1.1.1 and 1.2.2). Although acyl migration would likely be promoted in the employed hydrophilic solvent, the observed regioisomeric distributions, with the major products being the 2-, 3′- and 1′-regioisomers, largely conformed to the reactivity predicted by the electrostatic potential of the sucrose substituents, indicating that acyl migration was not favoured within the studied reaction time.

4 CONCLUSION AND OUTLOOK

The research presented in this work studied, developed and optimised quantifiable elution methods for RP-HPLC analysis of sucrose alkanoate regioisomers and investigated catalytic activity and regioselectivity in sucrose laurate syntheses in *N*,*N*-dimethylformamide with and without protease biocatalysts.

Design of experiments and statistical and multivariate data analysis were shown to be applicable and useful tools for method development in RP-HPLC analysis of sucrose alkanoates. The methodology was applied to the investigation, modelling and analysis of several elution strategies and the optimisation of sucrose alkanoate regioisomer separation. Although the elution method transferability between sucrose alkanoates of different alkanoyl chain lengths was shown to be limited, the application of DOE methodology would facilitate the adaption of elution methods for new analytes.

The aggregate objective function *resolution deviation* (RD) was defined for sample-specific evaluation of the achieved separation of three specific regioisomers, 3-, 6- and 1′-, showing similar elution properties in the RP-HPLC analysis of sucrose caprate. The RD was further developed into the sample- and system-independent resolution metric *general resolution deviation for multiple peaks* (RD_m) for evaluation of overall separation of any number of peaks of varying areas in chromatographic analysis.

The efficiency of different elution strategies was evaluated in terms of RD_m and analysis time. The best results for sucrose caprate were offered by step-down gradient elution, while isocratic elution with increased flow offered the best results for sucrose laurate. Step-down gradient elution offered improvements in separation by 13-38 % and reductions in analysis time by 3-9 % compared to isocratic elution for sucrose caprate. For sucrose laurate, isocratic elution with increased flow showed improvements in separation of up to 29 % and reductions of 6-34 % in analysis time compared to isocratic elution. The optimised elution strategies for the sucrose alkanoates resulted in the separation of regioisomers with resolutions adequate for quantification. All regioisomers of the sucrose alkanoates, as separated by RP-HPLC, were successfully identified and assigned.

The acylation of sucrose with vinyl laurate acyl donor in the hydrophilic, aprotic solvent *N*,*N*-dimethylformamide with or without protease biocatalyst resulted in the formation of sucrose laurate, with 2-*O*-lauroyl sucrose the most abundant monoester regioisomer synthesised. The serine protease ALP-901 produced the highest yield after 48 hours reaction time (53 %), while the reaction with thermolysin achieved the overall highest yield (63 %) after 6 hours, with only monoesters synthesised. The highest concentration of 2-*O*-lauroyl sucrose observed was 23.7 mM after 24 hours in the thermolysin-catalysed reaction. The detected appearance of the sucrose laurate regioisomers largely corresponded to the apparent rates of formation, based on monoester concentration and regioisomeric distribution after 48 hours, and 2-*O*-lauroyl sucrose was among the first regioisomers to appear in all investigated reactions. The concentration of sucrose laurate was observed to decline with reaction times above 6 hours in the reaction with thermolysin due to oligoester formation, and declining concentration with increasing reaction time was also observed for some specific sucrose laurate regioisomers in the reactions with

other enzyme formulations. Differences in regioisomeric distribution after 48 hours appeared partly to stem from differences in overall reaction rates. The regioisomeric distribution observed in the reaction without protein after 6 hours was similar to the distributions in the reactions with alcalase and casein after 48 hours. In addition, the distribution after 6 hours in the reaction with thermolysin was similar to the regioisomeric distribution in the reaction with ALP-901 after 48 hours, and similar degrees of sucrose conversion were observed for these reactions at the respective reaction times.

The reaction between sucrose and vinyl laurate with no protein in the reaction mixture appeared to be catalysed by the presence of aluminosilicate molecular sieves in the reaction medium. Non-catalytic protein in the reaction medium seemed to lower the catalytic activity of the molecular sieves. The catalytic functionality of aluminosilicate molecular sieves and their possible interaction effect on enzyme catalytic activity warrants further study, as the results presented here were rather limited.

MALDI mass spectrometry of 6-*O*-lauroyl sucrose and 6′-*O*-lauroyl sucrose resulted in signals characteristic for the molecular ion adducts with sodium and potassium ($[M+Na]^+$ and $[M+K]^+$, *m*/*z* 547 and 563, respectively). Tandem MS on the precursor ions resulted in different product ion mass spectra for the two regioisomers. These results indicated difference in stability of the molecular ion and fragmentation ion species between the regioisomers. The scope of the investigation was too narrow to determine whether MS/MS product ion mass spectra can be used for identification of the specific regioisomers, or if the extent of difference discernable by MS is whether the ester substitution is on the glucose or the fructose residue. Further study of sucrose laurate regioisomers acylated at secondary hydroxy positions and on the same saccharide residue analysed by mass spectrometry is necessary to draw firm conclusions.

BIBLIOGRAPHY

Allen, D. K.; Tao, B. Y. 1999. Carbohydrate-alkyl ester derivatives as biosurfactants. *J. Surfactants Deterg., 2* (3), 383-390.

Almarsson, O.; Klibanov, A. M. 1996. Remarkable activation of enzymes in nonaqueous media by denaturing organic cosolvents. *Biotechnol. Bioeng., 49* (1), 87-92.

Baker, I.; Matthews, B.; Suares, H.; Krodkiewska, I.; Furlong, D.; Grieser, F., et al. 2000. Sugar fatty acid ester surfactants: Structure and ultimate aerobic biodegradability. *J. Surfactants Deterg., 3* (1), 1-11.

Borges, M. R.; Balaban, R. 2007. Sucrose-branched polymer synthesized by protease from *Bacillus subtilis*. *Macromol. Symp., 258*, 25-29.

Christiansen, A.; Backensfeld, T.; Kuehn, S.; Weitschies, W. 2011. Investigating the stability of the nonionic surfactants tocopheryl polyethylene glycol succinate and sucrose laurate by HPLC-MS, DAD, and CAD. *J. Pharm. Sci., 100* (5), 1773-1782.

Christofides, J. C.; Davies, D. B. 1985. Co-operative and competitive hydrogen-bonding in sucrose determined by Simple-^{1}H-NMR-spectroscopy. *J. Chem. Soc. , Chem. Commun.* (21), 1533-1534.

Cruces, M.; Plou, F.; Ferrer, M.; Bernabe, M.; Ballesteros, A. 2001. Improved synthesis of sucrose fatty acid monoesters. *J. Am. Oil Chem. Soc., 78* (5), 541-546.

Dean, A. M.; Voss, D. 1998. *Design & Analysis of Experiments*. Springer: Secaucus, NJ, USA.

Degn, P.; Zimmermann, W. 2001. Optimization of carbohydrate fatty acid ester synthesis in organic media by a lipase from Candida antarctica. *Biotechnol. Bioeng., 74* (6), 483-491.

Devulapalle, K. S.; Gómez de Segura, A.; Ferrer, M.; Alcalde, M.; Mooser, G.; Plou, F. J. 2004. Effect of carbohydrate fatty acid esters on *Streptococcus sobrinus* and glucosyltransferase activity. *Carbohydr. Res., 339* (6), 1029-1034.

Dixon, R. W.; Peterson, D. S. 2002. Development and testing of a detection method for liquid chromatography based on aerosol charging. *Anal. Chem., 74* (13), 2930-2937.

Esbensen, K. H.; Guyot, D.; Westad, F. 2002. *Multivariate Data Analysis in Practice*, 5th ed. CAMO Software AS: Oslo.

Faber, K. 2004. *Biotransformations in Organic Chemistry: A Textbook*, 5th ed. Springer: Heidelberg.

Ferrer, M.; Soliveri, J.; Plou, F. J.; Lopez-Cortes, N.; Reyes-Duarte, D.; Christensen, M., et al. 2005. Synthesis of sugar esters in solvent mixtures by lipases from *Thermomyces lanuginosus* and *Candida antarctica* B, and their antimicrobial properties. *Enzyme Microb. Technol., 36* (4), 391-398.

Ferrer, M.; Cruces, M. A.; Bernabe, M.; Ballesteros, A.; Plou, F. J. 1999. Lipase-catalyzed regioselective acylation of sucrose in two-solvent mixtures. *Biotechnol. Bioeng., 65* (1), 10-16.

Fitzpatrick, P. A.; Klibanov, A. M. 1991. How can the solvent affect enzyme enantioselectivity. *J. Am. Chem. Soc., 113* (8), 3166-3171.

Fox, C. B.; Sivananthan, S. J.; Mikasa, T. J. T.; Lin, S.; Parker, S. C. 2013. Charged aerosol detection to characterize components of dispersed-phase formulations. *Adv. Colloid Interface Sci., 199–200* (0), 59-65.

Friebolin, H. 2005. *Basic One- and Two-Dimensional NMR Spectroscopy*. WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim.

García-Álvarez-Coque, M. C.; Torres-Lapasió, J. R.; Baeza-Baeza, J. J. 2006. Models and objective functions for the optimisation of selectivity in reversed-phase liquid chromatography. *Anal. Chim. Acta, 579* (2), 125-145.

Garti, N.; Clement, V.; Leser, M.; Aserin, A.; Fanun, M. 1999. Sucrose ester microemulsions. *J. Mol. Liq., 80* (2-3), 253-296.

Gupta, R. K.; James, K.; Smith, F. J. 1983. Sucrose esters and sucrose ester glyceride blends as emulsifiers. *J. Am. Oil Chem. Soc., 60* (4), 862-869.

Habulin, M.; Sabeder, S.; Knez, Z. 2008. Enzymatic synthesis of sugar fatty acid esters in organic solvent and in supercritical carbon dioxide and their antimicrobial activity. *J. Supercrit. Fluids, 45* (3), 338-345.

Huang, D.; Jiang, X.; Zhu, H.; Fu, X.; Zhong, K.; Gao, W. 2010. Improved synthesis of sucrose fatty acid monoesters under ultrasonic irradiation. *Ultrason. Sonochem., 17* (2), 352-355.

Hutchinson, J. P.; Li, J.; Farrell, W.; Groeber, E.; Szucs, R.; Dicinoski, G., et al. 2010. Universal response model for a corona charged aerosol detector. *J. Chromatogr. A, 1217* (47), 7418-7427.

Kato, A.; Arima, K. 1971. Inhibitory effect of sucrose ester of lauric acid on growth of *Escherichia coli*. *Biochem. Biophys. Res. Commun., 42* (4), 596-601.

Kato, S.; Kobayashi, H.; Watanabe, T. 1987. Antimicrobial action of sucrose monolaurate on *Staphylococcus aureus*. *J. Food Hyg. Soc. Japan, 28* (4), 261-266.

Koeller, K.; Wong, C. 2001. Enzymes for chemical synthesis. *Nature, 409* (6817), 232-240.

Kromidas, S. 2006. Fundamentals of Optimization: Principles of the Optimization of HPLC Illustrated by RP-Chromatography. In *HPLC Made to Measure: A Practical Handbook for Optimization*; Kromidas, S., Ed. Wiley-VCH Verlag GmbH & Co. KGaA: pp 3.

Lichtenthaler, F. W.; Immel, S.; Pokinskyj, P. 1995. Molecular Modeling of Saccharides .8. Selective 2-O-Benzylation of Sucrose - a Facile Entry to its 2-Deoxy-Derivatives and 2-Keto-Derivatives and to Sucrosamine. *Liebigs Annalen* (11), 1939-1947.

Lisa, M.; Lynen, F.; Holcapek, M.; Sandra, P. 2007. Quantitation of triacylglycerols from plant oils using charged aerosol detection with gradient compensation. *J. Chromatogr. A, 1176* (1-2), 135-142.

Moh, M. H.; Tang, T. S.; Tan, G. H. 2000. Improved separation of sucrose ester isomers using gradient high performance liquid chromatography with evaporative light scattering detection. *Food Chem., 69* (1), 105-110.

Molinier, V.; Wisniewski, K.; Bouchu, A.; Fitremann, J.; Queneau, Y. 2003. Transesterification of Sucrose in Organic Medium: Study of Acyl Group Migrations. *J. Carbohydr. Chem., 22* (7-8), 657-669.

Monk, J. D.; Beuchat, L. R.; Hathcox, A. K. 1996. Inhibitory effects of sucrose monolaurate, alone and in combination with organic acids, on *Listeria monocytogenes* and *Staphylococcus aureus*. *J. Appl. Bacteriol., 81* (1), 7-18.

Montgomery, D. C. 2012. *Design and Analysis of Experiments*, International Student Version, 8th ed. John Wiley & Sons: Singapore.

Nic, M.; Jirat, J.; Kosata, B. 2012. *IUPAC Compendium of Chemical Terminology (Gold Book).* http://goldbook.iupac.org/index.html (accessed JAN 10, 2014).

Ogino, H.; Ishikawa, H. 2001. Enzymes which are stable in the presence of organic solvents. *J. Biosci. Bioeng., 91* (2), 109-116.

Okabe, S.; Suganuma, M.; Tada, Y.; Ochiai, Y.; Sueoka, E.; Kohya, H., et al. 1999. Disaccharide esters screened for inhibition of tumor necrosis factor-alpha release are new anti-cancer agents. *Jap. J. Cancer Res., 90* (6), 669-676.

Okumura, H.; Kitazawa, N.; Wada, S.; Hotta, H. 2011. Stability of sucrose fatty acid esters under acidic and basic conditions. *J. Oleo Sci., 60* (6), 313-320.

Ottoson, J. 2001. Enthalpy and entropy in enzyme catalysis - A study of lipase enantioselectivity. Department of Biotechnology, Royal Institute of Technology , KTH, Stockholm, Sweden, Stockholm, Sweden.

Patil, D. R.; Rethwisch, D. G.; Dordick, J. S. 1991. Enzymatic synthesis of a sucrose-containing linear polyester in nearly anhydrous organic media. *Biotechnol. Bioeng., 37* (7), 639-646.

Pedersen, N. R.; Halling, P. J.; Pedersen, L. H.; Wimmer, R.; Matthiesen, R.; Veltman, O. R. 2002a. Efficient transesterification of sucrose catalysed by the metalloprotease thermolysin in dimethylsulfoxide. *FEBS Letters, 519* (1-3), 181-184.

Pedersen, N. R.; Wimmer, R.; Emmersen, J.; Degn, P.; Pedersen, L. H. 2002b. Effect of fatty acid chain length on initial reaction rates and regioselectivity of lipase-catalysed esterification of disaccharides. *Carbohydr. Res., 337* (13), 1179-1184.

Pedersen, N. R.; Wimmer, R.; Matthiesen, R.; Pedersen, L. H.; Gessesse, A. 2003. Synthesis of sucrose laurate using a new alkaline protease. *Tetrahedron: Asymmetry, 14* (6), 667-673.

Perez-Victoria, I.; Morales, J. C. 2006. Regioselectivity in acylation of oligosaccharides catalyzed by the metalloprotease thermolysin. *Tetrahedron, 62* (10), 2361-2369.

Perez-Victoria, I.; Zafra, A.; Morales, J. C. 2007. Determination of regioisomeric distribution in carbohydrate fatty acid monoesters by LC-ESI-MS. *Carbohydr. Res., 342* (2), 236-242.

Piao, J.; Kawahara-Aoyama, Y.; Inoue, T.; Adachi, S. 2006. Bacteriostatic activities of monoacyl sugar alcohols against thermophilic sporeformers. *Biosci. Biotech. Biochem., 70* (1), 263-265.

Plou, F. J.; Cruces, M. A.; Ferrer, M.; Fuentes, G.; Pastor, E.; Bernabe, M., et al. 2002. Enzymatic acylation of di- and trisaccharides with fatty acids: Choosing the appropriate enzyme, support and solvent. *J. Biotechnol., 96* (1), 55-66.

Plou, F. J.; Cruces, M. A.; Pastor, E.; Ferrer, M.; Bernabe, M.; Ballesterose, A. 1999. Acylation of sucrose with vinyl esters using immobilized hydrolases: Demonstration that chemical catalysis may interfere with enzymatic catalysis. *Biotechnol. Lett., 21* (7), 635-639.

Polat, T.; Linhardt, R. J. 2001. Syntheses and applications of sucrose-based esters. *J. Surfactants Deterg., 4* (4), 415-421.

Potier, P.; Bouchu, A.; Gagnaire, J.; Queneau, Y. 2001. Proteinase N-catalysed regioselective esterification of sucrose and other mono- and disaccharides. *Tetrahedron: Asymmetry, 12* (17), 2409-2419.

Puterka, G. J.; Farone, W.; Palmer, T.; Barrington, A. 2003. Structure-function relationships affecting the insecticidal and miticidal activity of sugar esters. *J. Econ. Entomol., 96* (3), 636-644.

Queneau, Y.; Fitremann, J.; Trombotto, S. 2004. The chemistry of unprotected sucrose: The selectivity issue. *C. R. Chimie, 7* (2), 177-188.

Queneau, Y.; Gagnaire, J.; West, J.; Mackenzie, G.; Goodby, J. 2001. The effect of molecular shape on the liquid crystal properties of the mono-O-(2-hydroxydodecyl)sucroses. *J. Mater. Chem., 11* (11), 2839-2844.

Queneau, Y.; Chambert, S.; Besset, C.; Cheaib, R. 2008. Recent progress in the synthesis of carbohydrate-based amphiphilic materials: the examples of sucrose and isomaltulose. *Carbohydr. Res., 343* (12), 1999-2009.

Reyes-Duarte, D.; Lopez-Cortes, N.; Ferrer, M.; Plou, F. J.; Ballesteros, A. 2005. Parameters affecting productivity in the lipase-catalysed synthesis of sucrose palmitate. *Biocatal. Biotransform., 23* (1), 19-27.

Ritthitham, S.; Wimmer, R.; Stensballe, A.; Pedersen, L. H. 2009a. Analysis and purification of O-decanoyl sucrose regio-isomers by reversed phase high pressure liquid chromatography with evaporative light scattering detection. *J. Chromatogr. A, 1216* (25), 4963-4967.

Ritthitham, S.; Wimmer, R.; Stensballe, A.; Pedersen, L. H. 2009b. Selectivity and stability of alkaline protease AL-89 in hydrophilic solvents. *J. Molec. Catal. B: Enzym., 59* (4), 266-273.

Riva, S.; Chopineau, J.; Kieboom, A. P. G.; Klibanov, A. M. 1988. Protease-catalyzed regioselective esterification of sugars and related compounds in anhydrous dimethylformamide. *J. Am. Chem. Soc., 110* (2), 584-589.

Robyt, J. F. 1998. *Essentials of carbohydrate chemistry*. Springer: New York.

Sadek, P. C. 2004. *Illustrated Pocket Dictionary of Chromatography*. Wiley: Hoboken, NJ, USA.

Shapiro, S. S.; Wilk, M. B. 1965. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika, 52* (3/4), 591-611.

Shi, Y. G.; Li, J. R.; Chu, Y. H. 2011. Enzyme-catalyzed regioselective synthesis of sucrose-based esters. *J. Chem. Technol. Biotechnol., 86* (12), 1457-1468.

Swartz, M. 2010. HPLC detectors: A brief review. *J. Liq. Chromatogr. Rel. Technol., 33* (9-12), 1130-1150.

Tamura, M.; Takahashi, A.; Uyama, A.; Mochizuki, N. 2010. Simultaneous analysis for carbohydrates in drinks by high performance liquid chromatography with charged aerosol detector. *Bunseki Kagaku, 59* (3), 239-245.

Taylor, V. F.; March, R. E.; Longerich, H. P.; Stadey, C. J. 2005. A mass spectrometric study of glucose, sucrose, and fructose using an inductively coupled plasma and electrospray ionization. *Int. J. Mass Spectrom., 243* (1), 71-84.

Thevenet, S.; Wernicke, A.; Belniak, S.; Descotes, G.; Bouchu, A.; Queneau, Y. 1999. Esterification of unprotected sucrose with acid chlorides in aqueous medium: Kinetic reactivity versus acyl- or alkyloxycarbonyl-group migrations. *Carbohydr. Res., 318* (1-4), 52-66.

Tijssen, R. 1998. The Mechanisms and Importance of Zone-Spreading. In *Handbook of HPLC*; Katz, E., Eksteen, R. and Schoenmakers, P., Eds. Marcel Dekker: New York, NY, USA, pp 55-142.

Torres, M. C.; Dean, M. A.; Wagner, F. W. 1990. Chromatographic separations of sucrose monostearate structural isomers. *J. Chromatogr., 522*, 245-253.

Ushikusa, T.; Maruyama, T.; Niiya, I.; Okada, M. 1990. Pyrolysis behavior and thermostability of sucrose ester of fatty acids. *J. Jap. Oil Chem. Soc., 39* (1), 38-41.

Vehovec, T.; Obreza, A. 2010. Review of operating principle and applications of the charged aerosol detector. *J. Chromatogr. A, 1217* (10), 1549-1556.

Vervoort, N.; Daemen, D.; Toeroek, G. 2008. Performance evaluation of evaporative light scattering detection and charged aerosol detection in reversed phase liquid chromatography. *J. Chromatogr. A, 1189* (1-2), 92-100.

Vulfson, E. N. 1993. Enzymatic synthesis of food ingredients in low-water media. *Trends Food Sci. Technol., 4* (7), 209-215.

Wang, C.; Li, J.; Yao, S.; Guo, Y.; Xia, X. 2007a. High-sensitivity matrix-assisted laser desorption/ionization Fourier transform mass spectrometry analyses of small carbohydrates and amino acids using oxidized carbon nanotubes prepared by chemical vapor deposition as matrix. *Anal. Chim. Acta, 604* (2), 158-164.

Wang, Q.; Zhang, S.; Yang, J. 2007b. HPLC analysis of sucrose ester analogs using evaporative light scattering detection. *J. Liq. Chromatogr. Rel. Technol., 30* (13-16), 2395-2407.

Wang, Q.; Zhang, S.; Zhang, P.; Zhu, J.; Yang, J. 2006. Separation and quantitation of sucrose esters using HPLC with evaporative light scattering detection. *J. Liq. Chromatogr. Rel. Technol., 29* (16), 2399-2412.

Watanabe, T.; Katayama, S.; Matsubara, M.; Honda, Y.; Kuwahara, M. 2000. Antibacterial carbohydrate monoesters suppressing cell growth of *Streptococcus mutans* in the presence of sucrose. *Curr. Microbiol., 41* (3), 210-213.

Weber, H. K.; Stecher, H.; Faber, K. 1995. Sensitivity of microbial lipases to acetaldehyde formed by acyl-transfer reactions from vinyl esters. *Biotechnol. Lett., 17* (8), 803-808.

Part II

CONTENTS OF PART II

1 DESIGN OF EXPERIMENTS AND MULTIVARIATE ANALYSIS FOR EVALUATION OF REVERSED-PHASE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY WITH CHARGED AEROSOL DETECTION OF SUCROSE CAPRATE REGIOISOMERS

Lie A., Wimmer R., and Pedersen L.H.

Journal of Chromatography A (2013), 1281, 67-72

Journal of Chromatography A, 1281 (2013) 67–72

Contents lists available at SciVerse ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Design of experiments and multivariate analysis for evaluation of reversed-phase high-performance liquid chromatography with charged aerosol detection of sucrose caprate regioisomers

Aleksander Lie, Reinhard Wimmer, Lars Haastrup Pedersen*

Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Sohngaardsholmsvej 49, DK 9000 Aalborg, Denmark

ARTICLE INFO

Article history:
Received 8 August 2012
Received in revised form 14 January 2013
Accepted 20 January 2013
Available online 25 January 2013

Keywords:
Sugar fatty acid ester regioisomers
Resolution deviation
RP-HPLC–CAD
Design of experiments
Elution efficiency

ABSTRACT

The use of step-down gradient elution profiles to improve separation of sucrose caprate regioisomers was investigated as part of the development of a quantitative RP-HPLC analysis method with charged aerosol detection. The investigation was conducted using design-of-experiments methodology and evaluated by multivariate regression analysis. This approach was proven to be useful for systematic method development in HPLC analysis. The gradient elution profiles were described by four variables – two concentration variables and two duration variables. The regression analysis showed that the concentration variables had the most significant effects on retention times, both as individual terms and as part of variable interactions. All the regioisomers exhibited non-linear relationships between eluent acetonitrile concentration and retention time with similar curvatures. Kendall rank correlation coefficients confirmed that the curvatures of the regioisomer curves were highly dependent on each other. Charged aerosol detection provided a mass-sensitivity of 10–100 ng for the sucrose fatty acid ester regioisomers. *Resolution deviation* (RD) was defined as an aggregate objective function for evaluating the separation of three specific sucrose caprate regioisomers with similar elution properties substituted at positions 6-, 3- and 1′-, respectively. The investigation resulted in the development of elution strategies for separation and quantitative RP-HPLC analysis of regioisomers of sucrose caprate with all eight sucrose caprate regioisomers successfully identified. Thus, resolutions above the level of adequacy for quantification, $R_s \geq 1.0$, were achieved for all regioisomers, both with isocratic and gradient elution strategies. For isocratic elutions, the best separation was achieved with eluent acetonitrile concentration 34%. Gradient elution resulted in a similar RD, but decreased the analysis time by 7–28%. For the gradient resulting in the most desirable combination of separation and analysis time, the R_s-values ranged from 1.31 to 6.82, and the analysis time was 24 min.

1. Introduction

Sugar fatty acid esters are non-ionic surfactants commonly applied as additives in foods, cosmetics, medical supplies and oral-care products [1,2]. Forming both water-in-oil and oil-in-water microemulsions [3], these compounds are used as emulsifiers and as essential ingredients in natural food aromas [4]. In particular, sugar monoesters of fatty acids, which have water solubility superior to the corresponding oligoesters, have been shown to possess antibiotic and insecticidal properties [1,5]. Fructose laurate has been shown to inhibit growth of *Streptococcus mutans* [6], and 6-*O*-lauroyl sucrose has been shown to inhibit growth of *Bacillus* sp. CECT 40 and *Lactobacillus plantarum* [1]. Thus, also being highly biodegradable and non-toxic to humans [7], sugar fatty acid monoesters make up a very interesting class of compounds for the development of new antimicrobials.

The physical and chemical properties of sugar fatty acid esters depend on the saccharide moiety, fatty acid chain length, and both position and degree of esterification [2]. Saccharides are polyols, giving rise to vast numbers of isomers. In the case of sucrose (see Fig. 1) esterification with a single fatty acid potentially results in a total of 255 isomers, eight of which are monoester regioisomers. Thus, sugar fatty acid esters can be synthesized in a wide range of hydrophilic-lipophilic balance ratios with different emulsifying and dispersing properties. The critical micelle concentration (CMC) of regioisomers of specific sucrose esters has been reported to vary significantly. The CMCs of 1′-*O*-myristoyl sucrose and 6-*O*-myristoyl sucrose were shown to differ by a factor of 1.4, while the CMC of 6-*O*-lauroyl sucrose was approximately 2.5 times greater than that of 1′-*O*-lauroyl sucrose [8].

* Corresponding author. Tel.: +45 9940 8497; fax: +45 9814 1808.
E-mail addresses: alie@bio.aau.dk (A. Lie), rw@bio.aau.dk (R. Wimmer), lhp@bio.aau.dk (L.H. Pedersen).

0021-9673/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.chroma.2013.01.079

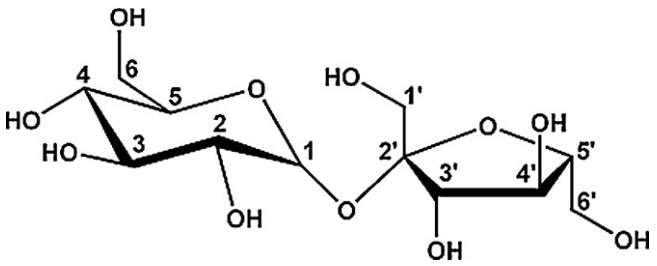


Fig. 1. Sucrose structure. Haworth perspective formula of sucrose with carbon atom numbering.

Separation of sucrose caprate regioisomers using RP-HPLC with C18-columns and acetonitrile-in-water eluents has been previously reported. Perez-Victoria et al. [9] reported the separation of all eight regioisomers by isocratic elution using RP-HPLC with electrospray ionization MS, without showing full separation of more than seven regioisomers. Ritthitham et al. [10] reported using preparative RP-HPLC with evaporative light-scattering detection to achieve separation of seven regioisomers of sucrose caprate by gradient elution, with a group of three regioisomers showing quite low resolution. They also reported that improved separation of the sucrose caprate regioisomers was obtained by decreasing the eluent acetonitrile concentration from the initial concentration, followed by an increasing gradient. However, resolution data were not presented or discussed in either study.

Charged aerosol detection (CAD) is suitable for analysis of non-volatile compounds, and not dependent on physicochemical properties, such as light absorption. Specifically for carbohydrates and fatty acids, CAD has been shown to be a high-sensitivity detection method [11,12]. Characteristic of CAD is that the detection method is based on the trapping of charged analyte particles and that its mass-response relations follow exponential functions with exponents typically below 1.0 [13,14].

The present work investigated the use of step-down gradient elution profiles to improve regioisomer separation as part of the development of a quantitative RP-HPLC analysis method for sucrose caprate regioisomers, using CAD. The investigation was conducted using design-of-experiments (DOE) methodology, entailing a systematic approach to such tasks. The elution profiles were described using several variables, and the resulting data were analysed using multivariate analysis and regression methods [15].

2. Materials and methods

2.1. Materials

All solvents were of HPLC grade. Acetonitrile, methanol (MeOH) and trifluoroacetic acid (TFA, reagent grade) were from Sigma–Aldrich. Water (H_2O) of Type I purity (ELGA PURELAB flex) was degassed by sonication for 20 min (Bransonic 2510E-MT, water bath). Sucrose caprate was a commercial sample (S-1266, ~95%) from Sigma Chemical Company.

2.2. Software

The Unscrambler X (10.1, CAMO Software) was used to develop the experimental design and perform the multivariate analyses of the results. MATLAB (R2011b, MathWorks) was used for the initial handling and collating of chromatographic data and calculation of second-order differences, Kendall rank correlation coefficients, resolutions and resolution deviations (see below).

Table 1
Design variables. Variable symbols, short descriptions and ranges.

Variable	Description	Range
A (%)	Initial acetonitrile concentration	30–40
B (min)	Duration of *A*	2–6
C (%)	Successive acetonitrile concentration	20–40
D (min)	Duration of *C*	1–5

2.3. Design of experiments

The HPLC analysis of sucrose caprate was investigated using a design-of-experiments approach. Based on the results from Ritthitham et al. [10], the gradient elution profile was defined by four variables, as summarised in Table 1, describing variations in elution parameters, followed by a constant eluent composition. An initial eluent concentration of acetonitrile, *A* (%), was maintained for a duration, *B* (min), after which the acetonitrile concentration was changed to a new level, *C* (%), which was maintained for a duration, *D* (min). Finally, the acetonitrile concentration was changed to a concentration level (35%) which was maintained for the remainder of the analysis time (see Fig. 2). The durations of the concentration transitions were kept constant (0.5 min). Variable ranges were determined from preliminary isocratic elutions and range testing experiments. The response variables represented the retention times for the eight possible regioisomers of sucrose caprate (see Fig. 1).

The chosen design was a face-centred composite (FCC) optimization design with four variables, giving a design space in the shape of a tesseract (4-cube). Each variable was investigated at three levels: low, mean and high. The samples consisted of three groups: *cube samples*, including all combinations of high and low levels of all variables; *axial samples*, combining low and high levels of a single variable at a time with mean levels of the remaining variables; and the *centre point*, which combined mean levels of all variables. The design was performed in two replicates and with seven samplings of the centre point, giving 25 different gradient elution profiles and 55 samples in total.

The dataset resulting from the designed experiments was analysed by building a partial least-squares (PLS) regression model (using the NIPALS-algorithm) incorporating variable interaction effects between up to four design variables and exponential terms up to exponent 4. The significance of variable effects in the PLS-model was evaluated based on the size of the weighted regression coefficients and the results of cross-validation coefficient-variance significance testing, and in an iterative process insignificant variable effects were removed and the PLS-model recalculated, until only significant effects remained.

2.4. RP-HPLC analysis and chromatographic conditions

Sucrose caprate regioisomers (0.4 g/L in MeOH, 10 μL injections) were analysed using an HP1100 system (Agilent), with a Corona CAD charged aerosol detector (ESA, N_2 at 35.0 psi, nebulizer at ambient temperature). Stationary phase was a Symmetry C18 column (250 mm × 4.6 mm, 5 μm particle size, Waters) maintained at 45 °C. Eluent consisted of mixtures of water and acetonitrile, both with 0.05% (v/v) TFA, and the flow rate was 1.0 mL/min. The chromatographic data were collected using ChemStation software.

Isocratic elutions were performed for integer increments in the range 30–40% acetonitrile in water. Design sample gradient elutions were performed according to the elution profiles specified by the experimental design.

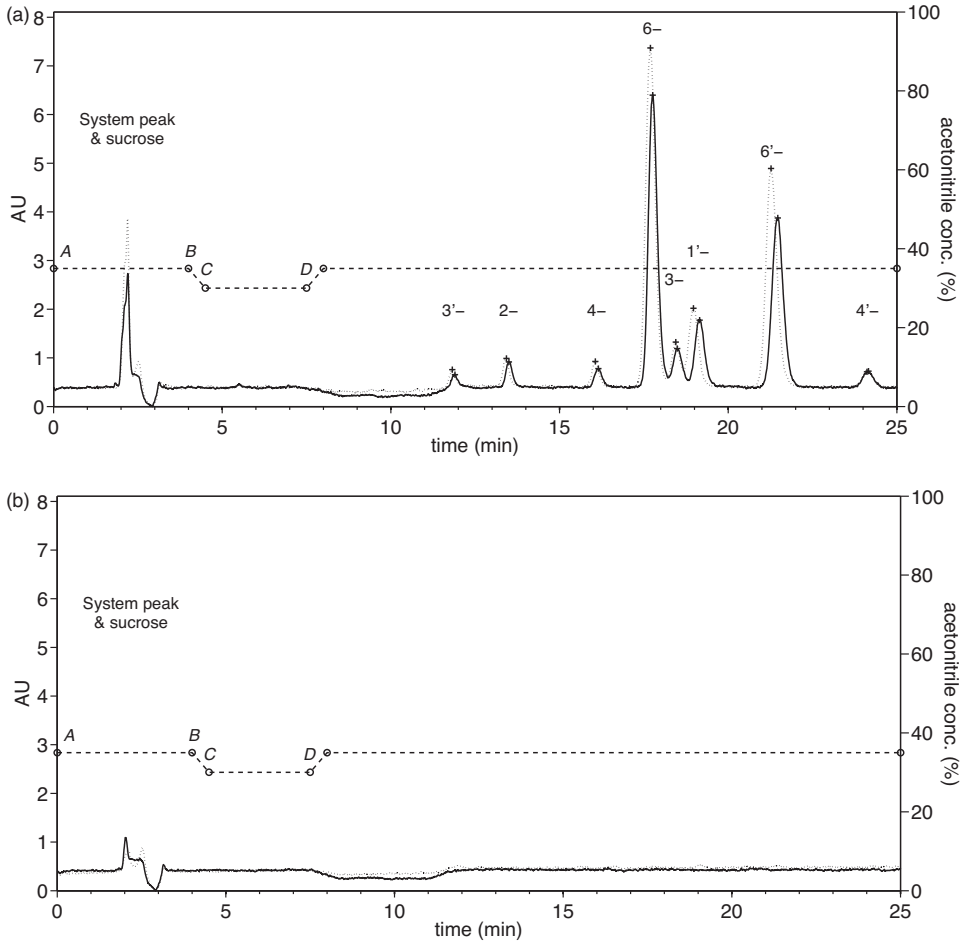


Fig. 2. (a) Chromatograms from RP-HPLC analysis of sucrose caprate, with profile of eluent acetonitrile concentration (dashed line). Elutions with TFA (solid line) and without TFA (dotted line) added to the eluents are shown. Variables *A*, *B*, *C* and *D* describe the elution profile: *A* = 35%, *B* = 4 min, *C* = 30%, *D* = 3 min. R_s-values: 3′- N/A; 2- 4.85; 4- 6.82; 6- 3.59; 3- 1.36; 1′- 1.31; 6′- 4.11; 4′- 4.30. (b) Chromatograms from RP-HPLC analysis of sample solvent (MeOH), with profile of eluent acetonitrile concentration (dashed line). Elutions with TFA (solid line) and without TFA (dotted line) added to the eluents are shown.

2.5. Regioisomer retention time as function of eluent acetonitrile concentration

For each sucrose caprate regioisomer, the curvature of the retention time data series was approximated using the second-order differences between the data points. A statistical comparison of the curvatures was performed by calculating the pairwise Kendall rank correlation coefficients.

2.6. Evaluation of separation: resolution deviation

Resolutions, expressed as R_s-values, were calculated from retention time and peak width according to Tijssen [16]. The two R_s-values for the separation of regioisomers 6-, 3- and 1′- (as annotated in Fig. 2a) were evaluated using an aggregate objective function (AOF) derived from root-mean-square deviation (RMSD), denominated *resolution deviation* (RD).

RMSD – a measure of the accuracy of predicted values – equals the square root of the variance, or the standard deviation, for unbiased estimators. The resolutions assessed here can be considered unbiased and, because only two R_s-values were compared, the sample standard deviation was applied:

$$SD(x) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (1)$$

where n is the number of observations, and $\bar{x}$ is the sample mean or the target value for comparison.

Baseline separation between chromatographic bands is obtained when $R_s = 1.5$ [16], which represents the optimal separation result, and any further increases in R_s would be neither beneficial nor detrimental. The two peak resolutions, $n = 2$, were constrained accordingly, $x_i = 0 \le R_{si} \le 1.5$, and optimal resolution was set as the comparison target, $\bar{x} = R_s = 1.5$. In order to facilitate comparison of RD-values, the AOF was scaled to an output range of [0, 1] (by multiplying with $a = 1/\sqrt{2 \times 1.5^2}$). From Eq. (1), the RD was defined as:

$$RD(R_{s1}-R_{s2}) = a \times \begin{cases} \sqrt{(R_{s1}-1.5)^2 + (R_{s2}-1.5)^2} & \text{if } R_{s1} < 1.5 \wedge R_{s2} < 1.5 \\ \sqrt{(R_{s1}-1.5)^2} & \text{if } R_{s1} < 1.5 \wedge R_{s2} \ge 1.5 \\ \sqrt{(R_{s2}-1.5)^2} & \text{if } R_{s1} \ge 1.5 \wedge R_{s2} < 1.5 \\ 0 & \text{if } R_{s1} \ge 1.5 \wedge R_{s2} \ge 1.5 \end{cases} \quad (2)$$

2.7. NMR spectroscopy

The position of acylation of the sucrose caprate regioisomers was determined by ^{1}H and ^{13}C NMR. All NMR spectra were recorded at 25 °C on a BRUKER AVIII NMR spectrometer operating at a field strength of 14.1T, and equipped with a 1.7 mm RT-TXI probe. Lyophilized samples were dissolved in methanol-d_4 (30 μL, Cambridge Isotope Laboratories) containing 0.05% (v/v)

Table 2
Chemical shifts of sucrose and sucrose caprate regioisomers in methanol-d_4 at 25 °C. Shifts are referenced relative to internal TMS. The chemical shifts of the C2′ carbon atoms were not recorded.

Atom	Sucrose	3′-	2-	4-	6-	3-	1′-	6′-	4′-
H1	5.39	5.39	5.53	5.44	5.38	5.44	5.40	5.35	5.42
C1	93.6	93.4	90.9	93.3	93.2	93.6	94.1	93.3	93.9
H2	3.42	3.42	4.60	3.52	3.45	3.58	3.42	3.45	3.45
C2	73.3	73.2	74.4	73.2	72.9	71.4	73.0	73.0	73.2
H3	3.70	3.61	3.88	3.88	3.74	5.21	3.68	3.74	3.73
C3	74.7	75.0	72.1	72.4	74.3	76.7	74.5	74.5	74.6
H4	3.35	3.39	3.42	4.81	3.52	3.52	3.37	3.36	3.38
C4	71.4	71.3	71.6	72.3	71.4	69.4	71.3	71.3	71.3
H5	3.82	3.81	3.89	4.02	4.03	3.91	3.83	3.84	3.88
C5	74.4	74.6	74.2	72.2	71.7	74.2	74.4	74.1	74.4
H6	3.80	3.80	3.82	3.60	4.39	3.81	3.80	3.84	3.84
	3.71	3.74	3.71	3.49	4.18	3.73	3.71	3.73	3.74
C6	62.2	62.3	62.2	62.0	64.6	61.8	62.1	62.4	62.2
H1′	3.64	3.62	3.51	3.65	3.64	3.66	4.37	3.64	3.68
	3.60	3.54	3.38	3.61	3.59	3.61	4.13		3.61
C1′	64.1	65.3	63.4	64.1	63.9	64.0	63.8	63.7	63.6
H3′	4.09	5.34	4.17	4.10	4.12	4.11	4.06	4.11	4.35
C3′	79.4	79.7	77.5	79.2	79.0	79.2	78.7	78.9	77.4
H4′	4.02	4.27	4.01	4.01	3.99	4.03	4.04	4.03	5.23
C4′	75.8	73.8	75.3	75.7	75.7	75.8	74.9	76.7	78.3
H5′	3.76	3.88	3.74	3.76	3.80	3.78	3.73	3.94	3.91
C5′	83.8	84.3	83.8	83.8	83.5	83.8	83.7	80.4	82.4
H6′	3.75	3.78	3.76	3.76	3.80	3.77	3.76	4.39	3.78
					3.75			4.32	3.73
C6′	63.3	63.1	63.5	63.5	63.8	63.4	63.2	66.9	63.8

tetramethylsilane (TMS). Resonance assignment and identification of compounds was achieved by help of COSY and HSQC spectra.

3. Results

3.1. Separation of sucrose caprate regioisomers

The RP-HPLC analysis of sucrose caprate resulted in the resolution of the eight possible regioisomers with R_S-values ranging from 1.31 to 6.82 (see Fig. 2a). The peaks in the chromatogram were confirmed as sucrose caprate regioisomers by analysis of the sample solvent (see Fig. 2b), and the identification of the individual regioisomers was determined by NMR spectroscopy (see Table 2). CAD showed detection of sucrose caprate with a sensitivity of 10–100 ng.

3.2. Dependence of retention time on eluent acetonitrile concentration

Increasing the eluent acetonitrile concentration decreased the retention times for the sucrose caprate regioisomers (see Fig. 3). The 3-regioisomer was resolved only in the concentration range 31–37%. At 30% the 3-regioisomer co-eluted with the preceding 6-regioisomer, while at 38% and above, it co-eluted with the following 1′-regioisomer.

All the regioisomers exhibited similar, non-linear relationships between eluent acetonitrile concentration and retention time. Kendall rank correlation coefficients (see Table 3) confirmed that the curvatures of the regioisomer curves were highly dependent on each other.

3.3. Design-of-experiments analysis

The effects of the four elution profile variables on the resulting regioisomer retention times were shown to have varying significance (see Table 4). The concentration variables (*A* and *C*) had the most significant effects, as isolated variables, as exponential terms, and as part of their interaction (*AC*). In addition, the interaction effect of the concentration variables and the duration of the initial concentration (*ABC*) was highly significant. The time variables were only significant as part of interaction effects involving one or both of the concentration variables.

Among the design samples, six samples had seven responses (with the 3-regioisomer unresolved). These samples were the replicates of cube samples with high level for the variable sets *ABC*, *ACD*, and *ABCD*. Based on the seven samplings of the centre point, the relative experimental error was 1.6%. In the iterative processing of the PLS regression model of the design dataset, the order of eliminated variable effects were as follows: *BD*, *DD*, *DDD*, *DDDD*, *D*, *B*, *BB*, *BBB*, *BBBB*, *AD*, *ABD*, *BCD*, *CD*.

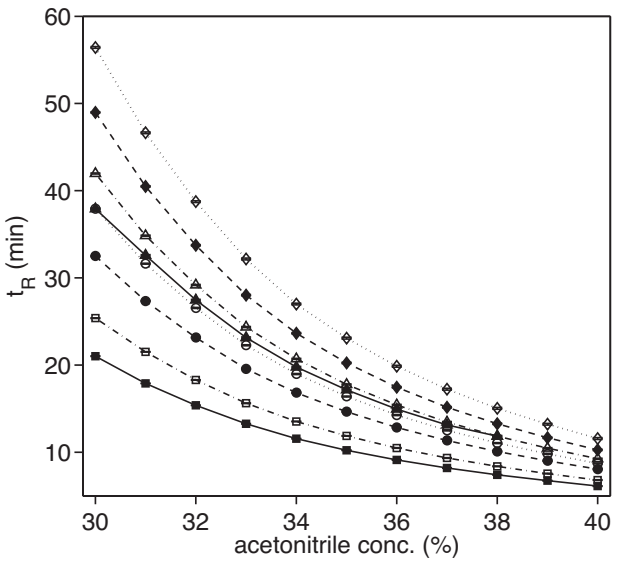


Fig. 3. Regioisomer retention times as function of eluent acetonitrile concentration for isocratic elutions. The regioisomers are arrayed from bottom to top according to elution order: 3′- solid line, filled square; 2- dash–dot line, open square; 4- dashed line, filled circle; 6- dotted line, open circle; 3- solid line, filled triangle; 1′- dash–dot line, open triangle; 6′- dashed line, filled diamond; 4′- dotted line, open diamond.

Table 3
Kendall rank correlation coefficients (τ) for the second-order differences of the isocratic elution data. For regioisomer pairs with correlation coefficient $\tau = 1$, all rank order pairs were concordant. For correlation coefficients of the 3-regioisomer ($\tau = 4/5$) there were 9 concordant pairs and 1 discordant pair. The curvatures were compared over as much as possible of the concentration range studied: 32–36% for the 3-regioisomer; 31–39% for the other regioisomers.

	3′-	2-	4-	6-	3-	1′-	6′-	4′-
3′-	*	1	1	1	4/5	1	1	1
2-		*	1	1	4/5	1	1	1
4-			*	1	4/5	1	1	1
6-				*	4/5	1	1	1
3-					*	4/5	4/5	4/5
1′-						*	1	1
6′-							*	1
4′-								*

Table 4
Significance of selected variable effects, from PLS analysis of face-centred composite design, using cross-validation coefficient-variance significance testing. Based on estimated p-values: +++ ($p < 0.005$), ++ ($p < 0.01$), + ($p <= 0.05$), 0 ($p > 0.05$). Effects not included in the table were insignificant ($p > 0.05$). The significance levels for the exponential terms (A^x, C^x) were equal for all exponents in the range $2 \leq x \leq 4$. Variables A–D are defined in Table 1 and their functions in the experimental design are outlined in Section 2.3.

Effect	Responses							
	3′-	2-	4-	6-	3-	1′-	6′-	4′-
A	+++	+++	+++	+++	+++	+++	+++	+++
B	0	0	0	0	0	0	0	0
C	+++	+++	+++	+++	+++	+++	+++	+++
D	0	0	0	0	0	0	0	0
AB	0	0	0	0	0	0	0	0
AC	+++	+++	+++	+++	+++	+++	+++	+++
BC	+	+	+	+	+	+	+	+
ABC	+++	+++	+++	+++	+++	+++	+++	+++
ACD	+	+	+	+	+	+	+	+
ABCD	+	+	+	+	+	+	0	0
A^x	+++	+++	+++	+++	+++	+++	+++	+++
C^x	+++	+++	+++	+++	+++	+++	+++	+++

3.4. Evaluation of elution efficiency

Among the isocratic elutions, the eluent concentration of 34% acetonitrile gave the lowest RD (see Table 5a), while the elutions with higher concentrations of acetonitrile showed shorter analysis time but higher RD.

With gradient elution the RD obtained was in the same range as for isocratic elution, however, the analysis time was shortened significantly. Thus, for RD-values in the range 0.15–0.16 a 28% reduction in analysis time was obtained, while for the RD-range 0.095–0.11 an average analysis time reduction of 7% was obtained (see Table 5b).

Table 5
Elution efficiency.

Acetonitrile conc. (%)	Resolution deviation	Analysis time (min)
(a) Isocratic elutions with all sucrose caprate regioisomer peaks resolved		
31	0.24 (3.8)	47
32	0.25 (4.4)	39
33	0.16 (3.9)	32
34	0.096 (73)	27
35	0.20 (1.7)	23
36	0.23 (1.4)	20
37	0.35 (2.5)	17

A (%), *B* (min), *C* (%), *D* (min)	Resolution deviation	Analysis time (min)
(b) Gradient elutions with all sucrose caprate regioisomer peaks resolved		
35, 4, 30, 3	0.11 (2.0)	24
35, 6, 30, 1	0.15 (2.0)	23
35, 2, 30, 1	0.15 (2.0)	23
30, 4, 30, 1	0.095 (20)	26

4. Discussion

The Kendall rank correlation analysis of the isocratic elutions showed a relationship between acetonitrile concentration and retention time with similar curvature for all the sucrose caprate regioisomers (see Fig. 3). Slight variations in curvature between regioisomers were caused by experimental variations, and the lower τ-value for one specific regioisomer (3-) was attributed to the lower number of observations and the narrow concentration range over which this particular regioisomer was studied. As the curvature of the retention time dependence on eluent composition was the same for all regioisomers, the differences between their dependence functions were expressed by a multiplicative factor and an additive offset specific for each regioisomer, reflecting differences in physicochemical properties causing the specific elution sequence and separation of the sucrose caprate regioisomers in the HPLC-analyses.

A gradient elution strategy was developed for quantification of all regioisomers of sucrose caprate. Baseline separation was achieved for the majority of the regioisomers (see Fig. 2a), while for three regioisomers (6-, 3- and 1′-) the separation obtained resulted in resolutions 9–13% lower than what was required for complete baseline separation, as shown by the R_s-values ($R_{s3} = 1.36$; $R_{s1'} = 1.31$). However, in most practical applications $R_s \geq 1.0$ is considered the level of adequacy for analytical quantification [16], and for the gradient elution profile resulting in the most desirable combination of separation and analysis time, R_s-values above this threshold were achieved (see Fig. 2a). Although isocratic elutions offered similar separations (in terms of RD-values), 7–28% shorter analysis time was achieved using the gradient elution strategy.

The developed AOF-expression, resolution deviation (RD), provided a simplified measure for evaluating the separation achieved by any elution strategy; however, some particular properties of the AOF should be pointed out. Firstly, the RD has a decreasing scale and approaches zero as the resolutions approach the target value indicating baseline separation. By applying the above mentioned level of adequacy for quantification to both component resolutions, the threshold of adequacy becomes RD ≤ 0.236. The isocratic elutions in the acetonitrile concentration range 33–36% and all gradient elutions listed in Table 5 exhibited RD-values below this threshold. Secondly, the RD only indicates the overall separation of the three regioisomers 6-, 3- and 1′-, and does not reveal anything about the relative sizes of the component resolutions. Thus, several combinations of R_s-values can result in the same RD.

A face-centred composite design conformed to the limitations of the variable space and the relationship between eluent composition and retention times. The non-linear relationship between eluent acetonitrile concentration and regioisomer retention times, as shown in this investigation, made the use of an optimization design necessary to model the design space adequately [15]. Also, because the 3-regioisomer was resolved only within a limited variable space, the operational variable space was strictly delimited and corresponded to the region of interest, causing severe challenges in the application of circular experimental designs.

In considering which variables effects were most influential on the regioisomer retention times, the eluent acetonitrile concentrations proved to have the highest significance, while a range of variable effects, including the individual terms for duration of maintaining specific eluent compositions, were determined as insignificant and could be disregarded. In particular, the concentration variables had significant effects, both as individual terms and as part of interactions. The duration variables were shown to

be significant only as part of interaction effects involving the concentration variables. The exponential terms of the concentration variables were also highly significant and were interpreted as an expression of the nonlinear relationships seen in Fig. 3.

5. Conclusion

Design of experiments and multivariate data analysis were shown to be applicable and useful tools for method development in RP-HPLC analysis of sucrose fatty acid esters. Using these tools, a method for quantitative RP-HPLC analysis of regioisomers of sucrose caprate was developed. The statistical analyses proved that the eluent acetonitrile concentrations had the highest significance, while other variable effects, including the duration of maintaining specific eluent compositions, were determined as insignificant. In particular, the concentration variables had significant effects both as individual terms and as part of interactions.

Resolutions above the level of adequacy for quantification, $R_S \geq 1.0$, were achieved for all regioisomers, both with isocratic and gradient elution strategies. For isocratic elutions, the best separation was achieved with eluent acetonitrile concentration 34%. Gradient elution resulted in a similar RD, but reduced the analysis time 7–28%. All the regioisomers of sucrose caprate were successfully identified and assigned and charged aerosol detection provided a mass-sensitivity of 10–100 ng.

RD was defined as an AOF for evaluating the separation of the three regioisomers substituted at the positions 6-, 3- and 1′-, respectively, showing similar elution properties.

Acknowledgement

The present research was funded by the Danish Council for Independent Research.

References

[1] M. Ferrer, J. Soliveri, F.J. Plou, N. Lopez-Cortes, D. Reyes-Duarte, M. Christensen, J.L. Copa-Patino, A. Ballesteros, Enzyme Microb. Technol. 36 (2005) 391.
[2] G.J. Puterka, W. Farone, T. Palmer, A. Barrington, J. Econ. Entomol. 96 (2003) 636.
[3] N. Garti, V. Clement, M. Leser, A. Aserin, M. Fanun, J. Mol. Liq. 80 (1999) 253.
[4] E.N. Vulfson, Trends Food Sci. Technol. 4 (1993) 209.
[5] F.J. Plou, M.A. Cruces, M. Ferrer, G. Fuentes, E. Pastor, M. Bernabe, M. Christensen, F. Comelles, J.L. Parra, A. Ballesteros, J. Biotechnol. 96 (2002) 55.
[6] T. Watanabe, S. Katayama, M. Matsubara, Y. Honda, M. Kuwahara, Curr. Microbiol. 41 (2000) 210.
[7] E. Castillo, F. Pezzotti, A. Navarro, A. Lopez-Munguia, J. Biotechnol. 102 (2003) 251.
[8] T. Polat, R.J. Linhardt, J. Surfactants Deterg. 4 (2001) 415.
[9] I. Perez-Victoria, A. Zafra, J.C. Morales, Carbohydr. Res. 342 (2007) 236.
[10] S. Ritthitham, R. Wimmer, A. Stensballe, L.H. Pedersen, J. Chromatogr. A 1216 (2009) 4963.
[11] T. Vehovec, A. Obreza, J. Chromatogr. A 1217 (2010) 1549.
[12] M. Tamura, A. Takahashi, A. Uyama, N. Mochizuki, Bunseki Kagaku 59 (2010) 239.
[13] N. Vervoort, D. Daemen, G. Toeroek, J. Chromatogr. A 1189 (2008) 92.
[14] J.P. Hutchinson, J. Li, W. Farrell, E. Groeber, R. Szucs, G. Dicinoski, P.R. Haddad, J. Chromatogr. A 1217 (2010) 7418.
[15] K.H. Esbensen, D. Guyot, F. Westad, Multivariate Data Analysis in Practice, CAMO Software AS, Oslo, 2002, pp. 598–xx.
[16] R. Tijssen, in: E. Katz, R. Eksteen, P. Schoenmakers (Eds.), The Mechanisms and Importance of Zone-Spreading. Handbook of HPLC, Marcel Dekker, New York, NY, USA, 1998, p. 55.

2 ELUTION STRATEGIES FOR REVERSED-PHASE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY ANALYSIS OF SUCROSE ALKANOATE REGIOISOMERS WITH CHARGED AEROSOL DETECTION

Lie A., and Pedersen L.H.

Journal of Chromatography A (2013), 1311, 127-133

Journal of Chromatography A, 1311 (2013) 127–133

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Elution strategies for reversed-phase high-performance liquid chromatography analysis of sucrose alkanoate regioisomers with charged aerosol detection

Aleksander Lie, Lars Haastrup Pedersen*

Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Sohngaardsholmsvej 49, 9000 Aalborg, Denmark

ARTICLE INFO

Article history:
Received 24 June 2013
Received in revised form 21 August 2013
Accepted 21 August 2013
Available online 27 August 2013

Keywords:
Sugar fatty acid ester regioisomers
Resolution deviation
RP-HPLC-CAD
Design of experiments
Elution efficiency

ABSTRACT

A broad range of elution strategies for RP-HPLC analysis of sucrose alkanoate regioisomers with CAD was systematically evaluated. The HPLC analyses were investigated using design-of-experiments methodology and analysed by analysis of variance (ANOVA) and regression modelling. Isocratic elutions, isocratic elutions with increased flow, and gradient elutions with step-down profiles and step-up profiles were performed and the chromatographic parameters of the different elution strategies were described by suitable variables. Based on peak resolutions general resolution deviation for multiple peaks (RD_m) was developed for sample-independent evaluation of separation of any number of peaks in chromatographic analysis. Isocratic elutions of sucrose alkanoates showed similar relationships between eluent acetonitrile concentration and retention time for all regioisomers of sucrose caprate and sucrose laurate, as confirmed by evaluation of the curvatures using approximate second derivatives and Kendall rank correlation coefficients. Regression modelling and statistical analysis showed that acetonitrile concentration and flow rate were highly significant for both average adjusted retention time and RD_m for sucrose laurate. For both responses the effect of changes in acetonitrile concentration was larger than the effect of changes in flow rate, over the ranges studied. Regression modelling of the step-down gradient profiles for the sucrose alkanoates showed that the eluent acetonitrile concentrations were the overall most significant variables for retention time and separation. The models for average adjusted retention time of sucrose caprate and sucrose laurate showed only a few differences in the significance levels of terms, while the models for RD_m showed larger differences between the sucrose alkanoates, in both the number of terms and their significance. Efficiency evaluation of elution strategies, in terms of RD_m and analysis time, showed that the best results were offered by step-down gradient elution for sucrose caprate and isocratic elution with increased flow for sucrose laurate. Step-down gradient elution of sucrose caprate offered improvements in separation at similar analysis time compared to isocratic elution, with the most efficient elutions achieved with elution profile acetonitrile concentrations at 32.5% and 25%, resulting in reduction of RD_m by 13–38% and reduction of analysis time by 3–9%. For sucrose laurate, isocratic elution with increased flow showed improvements in separation and reductions in analysis time compared to isocratic elution, such as elution at 37% with flow 2.0 mL/min resulting in reduction of analysis time by 34% and equal RD_m, while elution at 35% with flow 2.0 mL/min reduced RD_m by 29% and analysis time by 6%, compared to isocratic elution at 38% acetonitrile with standard flow (1.0 mL/min).

1. Introduction

Sugar fatty acid esters are non-ionic surfactants commonly applied as additives in foods, cosmetics, medical supplies and oral-care products [1,2]. These compounds are used as essential ingredients in natural food aromas and as emulsifiers, forming both water-in-oil and oil-in-water microemulsions [3,4]. In particular, sugar monoesters of fatty acids have shown interesting properties, such as water solubility superior to the corresponding oligoesters and antibiotic and insecticidal effects [1,5]. The physical and chemical properties of sugar fatty acid esters depend on the saccharide moiety, the fatty acid chain length, and the position and degree of esterification [2]. Thus, sugar fatty acid esters can be synthesised in a wide range of hydrophilic–lipophilic balance ratios with different emulsifying and dispersing properties. In the case of sucrose, esterification with a single fatty acid potentially results in a total of 255 isomers, eight of which are monoester regioisomers.

* Corresponding author. Tel.: +45 9940 8497; fax: +45 9814 1808.
E-mail addresses: alie@bio.aau.dk (A. Lie), lhp@bio.aau.dk (L.H. Pedersen).

0021-9673/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.chroma.2013.08.081

Table 1
Design variable symbols, definitions and ranges.

Variable	Description	Sucrose caprate	Sucrose laurate
A (%)	Initial acetonitrile concentration	30.0–40.0	38.0–46.0
B (min)	Duration of *A*	2.0–6.0	0.0–6.0
C (%)	Successive acetonitrile concentration	20.0–40.0	24.0–38.0
D (min)	Duration of *C*	1.0–5.0	2.0–6.0
F (mL/min)	Eluent flow	–	1.0–2.0
I (%)	Isocratic acetonitrile concentration	–	35.0–41.0

Analysis of sucrose alkanoate regioisomers by RP-HPLC with C18-columns and acetonitrile-in-water eluents has previously been investigated. Perez-Victoria et al. [6] reported the separation of all eight regioisomers of sucrose caprate, sucrose laurate and sucrose palmitate by isocratic elutions, without presenting resolution data. By gradient elution and preparative RP-HPLC, Ritthitham et al. [7] achieved separation of seven regioisomers of sucrose caprate, without baseline separation for a group of three regioisomers and resolutions were not discussed.

We have used statistical methods to investigate isocratic and step-down gradient elution strategies and achieved quantifiable separation of all eight regioisomers of sucrose caprate by RP-HPLC with charged aerosol detection (CAD) [8]. For evaluation of the separation of three specific sucrose caprate regioisomers with similar elution properties, a resolution-based aggregate objective function, resolution deviation (RD), was developed. Gradient elution profiles were described by two concentration variables and two duration variables and regression analysis showed that the concentration variables had the largest effects on retention times, as individual terms and as part of interactions with the duration variables. Step-down gradient elution was shown to offer shorter analysis time with similar separation as isocratic conditions, and CAD provided mass-sensitivity in the range of 10–100 ng for the sucrose caprate regioisomers.

In the present investigation a wide range of elution strategies for RP-HPLC analysis of sucrose alkanoate regioisomers with CAD was systematically evaluated. Design-of-experiments (DOE) methodology and statistical analysis were used to evaluate isocratic as well as complex gradient elution strategies, and the potential and limitations of methodology transfer for analysis of sucrose alkanoates of increasing alkanoyl chain length were investigated. A sample-independent *general resolution deviation for multiple peaks* (RD_m) was developed as a numerical measure to evaluate the separation obtained with any of the studied elution strategies.

2. Materials and methods

2.1. Materials

All solvents were of HPLC grade, unless otherwise indicated. Methanol (MeOH) and trifluoroacetic acid (TFA, reagent grade) were from Sigma–Aldrich, while acetonitrile (ACN) was from Panreac Quimica. Water (H_2O) of Type I purity (ELGA PURELAB flex) was degassed by sonication for 20 min (Bransonic 2510E-MT, water bath). Sucrose caprate was from Sigma–Aldrich (S-1266, ~95%), while sucrose laurate was from Carbosynth (dodecanoyl-D-sucrose).

2.2. Software

MATLAB (R2012b, The MathWorks, Inc.) was used for processing chromatographic data and calculation of resolutions and resolution deviations, as defined below. R (2.15.2, The R Foundation for Statistical Computing) was used for statistical analyses.

2.3. Design of experiments

The HPLC analyses of the sucrose alkanoates were investigated using design-of-experiments approaches with elution profiles described by variables as previously defined [8]. Based on the results from Ritthitham et al. [7], the gradient elution profile was composed of two eluent concentration variables (*A* and *C*, %) and two duration variables (*B* and *D*, min). Following the variables, eluent concentration was changed according to the fatty acid chain length (35% for sucrose caprate, 38% for sucrose laurate) and maintained for the remainder of the analysis time. In addition, for isocratic concentration and flow rate variations, two variables were investigated: eluent flow rate, *F* (mL/min), and isocratic acetonitrile concentration, *I* (%). The variables are summarised in Table 1, and their ranges were determined from preliminary experiments and HPLC-system specifications.

The responses analysed were average adjusted retention time ($\bar{t}'_R$, min), analysis time (min) and general resolution deviation (RD_m), as defined below. Distributions of responses were investigated using normal probability plots and the Shapiro–Wilk normality test [9]. If required, a transformation parameter was estimated using the Box–Cox method [10]. Experimental data were assessed using percent relative standard deviation (%RSD), and analysed by analysis of variance (ANOVA) and regression modelling on coded variables with the maximum number of exponential and interaction terms. The models were iteratively reduced, according to the hierarchy principle, through elimination of the least significant, highest-order term, until only significant terms remained ($\alpha = 0.1$), and evaluated based on distribution and variance of residuals, coefficients of determination (R^2 and $R^2_{adj.}$), and percent relative root-mean-square error (%RMSE).

Flow rate effects under isocratic conditions were investigated using a 4-level factorial design with added centre point, giving a square grid design, with each variable investigated at five levels. The design was performed in one replicate and with 4 samplings of the centre point, giving 17 unique HPLC analyses of 20 in total.

Gradient elution profiles were investigated using a circumscribed central composite (CCC) optimisation design with four variables, giving a 4-ball design space, with each variable investigated at five levels. The design was performed in one replicate and with six samplings of the centre point, giving 25 unique HPLC analyses of 30 in total.

2.4. Chromatographic conditions

Sucrose alkanoate regioisomers were analysed using RP-HPLC according to Lie et al. [8]. Samples of sucrose caprate (0.4 g/L in MeOH, 10 μL injections) and sucrose laurate (1.0 g/L in H_2O/ACN 1:1) were analysed using an HP1100 system (Agilent) with a Corona CAD charged aerosol detector (ESA, N_2 at 35.0 psi, nebuliser at ambient temperature). Stationary phase was a Symmetry C18 column (250 mm × 4.6 mm, 5 μm particle size, Waters) maintained at 45 °C. Eluent consisted of mixtures of water and acetonitrile, both with 0.05% (v/v) TFA, and the flow rate was 1.0 mL/min, unless

specified by the experimental design. The chromatographic data were collected using ChemStation software.

Isocratic elutions were performed in integer increments in the range 36–46% acetonitrile in water, and with acetonitrile concentrations and flow rates specified by the experimental design. Gradient elutions were performed with step-down profiles specified by the experimental design, and with step-up profiles specified by modification of step-down profiles by exchanging values for the variables A and C and changing the variables B and D to the respective mean values.

2.5. Retention time as function of acetonitrile concentration

The curvatures of the relationships between regioisomer retention time and eluent acetonitrile concentration were approximated and compared, using Kendall rank correlation coefficients (τ), as previously described [8].

2.6. General resolution deviation for multiple peaks

Based on peak resolutions, a generalised sample-independent term was developed for evaluation of the separation of any number of peaks, expanding a similar aggregate objective function previously defined for evaluating the separation of three specific sucrose caprate regioisomers [8]. This *general resolution deviation for multiple peaks* (RD_m) was derived on the following principles:

Tijssen [11] defines peak width as a function of retention time,

$$\sigma = \bar{\omega} t'_R \tag{1}$$

where $\bar{\omega}$ is the slope of the peak width and t'_R is the adjusted retention time, and peak width at baseline as

$$w_b = 4\sigma \tag{2}$$

where σ is the standard deviation. Eqs. (1) and (2) were combined to calculate $\bar{\omega}$ from the peak with the largest (maximum) area,

$$\bar{\omega} = \frac{w_b}{4t'_R} \quad \text{for } \max(A) \tag{3}$$

and this value was used, for all peaks of interest, to estimate peak widths representing sample components present in equal amounts.

Assuming that pairs of peaks with similar properties were closely eluted, the equation for the resolution [11], R_s, was combined with Eqs. (1) and (2), and simplified to

$$R_s = \frac{2(t_2 - t_1)}{w_1 + w_2} = \frac{2\Delta t_R}{4\sigma + 4\sigma} = \frac{\Delta t_R}{4\sigma} = \frac{\Delta t_R}{4\bar{\omega} t'_R} \tag{4}$$

Sample standard derivation is defined as

$$SD(x) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{5}$$

where n is the number of observations, and $\bar{x}$ is the sample mean or the target value for comparison.

Evaluating resolutions gave $x_i = R_s$, and the threshold target value for optimal resolution is given by $R_s \geq 1.5 = \bar{x}$. Increasing R_s beyond $R_s = 1.5$ does not represent any improvement in resolution, and this constraint was defined as a minimum function

$$\min(x_i - 1.5, 0) = \begin{cases} 0 & x_i > 1.5 \\ x_i - 1.5 & x_i \leq 1.5 \end{cases} \tag{6}$$

For any pair of peaks, (t_{i-1}, t_i), Eq. (4) gave the estimated resolution for the components present in equal amounts as

$$x_i = \frac{\Delta t_i}{4\sigma_i} = \frac{(t_i - t_{i-1})}{4\bar{\omega} t'_i} = \frac{(t_i - t_{i-1})}{4\bar{\omega}(t_i - t_0)} \tag{7}$$

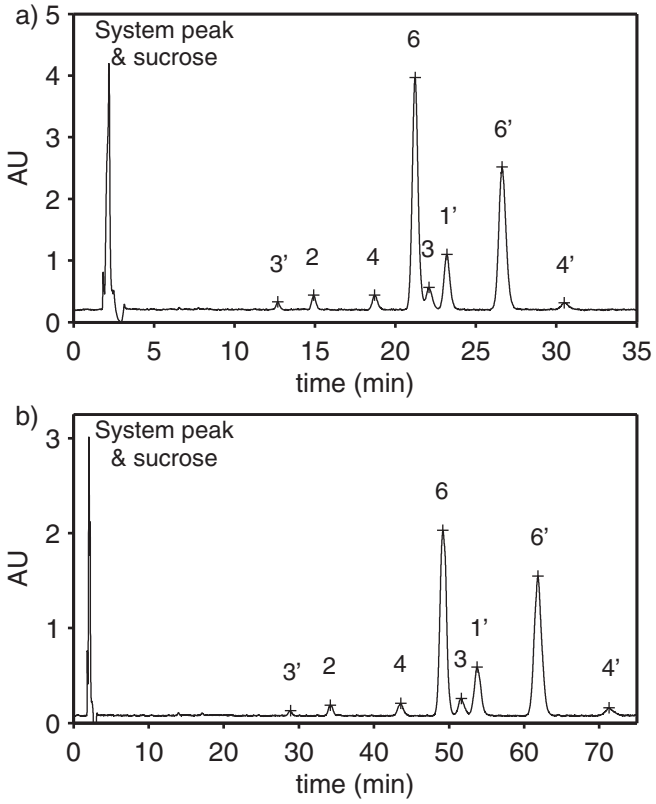


Fig. 1. Chromatograms from RP-HPLC analysis of sucrose alkanoates. (a) Sucrose caprate analysed with isocratic elution at 33% acetonitrile concentration. (b) Sucrose laurate analysed with isocratic elution at 36% acetonitrile concentration. Peaks are annotated with regioisomer substitution position.

where $i \in [2, n]$, as the number of calculated resolutions is $n - 1$, or 1 less than the number of peaks evaluated. Inserting Eqs. (6) and (7) in Eq. (5), gave the final expression for RD_m

$$RD_m(t_R) = a \times \sqrt{\frac{1}{n-1}\sum_{i=2}^{n} \min\left(\frac{(t_i - t_{i-1})}{4\bar{\omega}(t_i - t_0)} - 1.50\right)^2} \tag{8}$$

where a is a scaling constant. Scaling by $a = 1/1.5$ resulted in a range between 0 and 1, $RD_m \in [0, 1]$, with 1 representing no resolution and 0 representing optimal resolution for all evaluated peaks, as $RD_m \propto 1/\Sigma R_s$.

3. Results

3.1. Isocratic elutions

The RP-HPLC analyses of sucrose caprate and sucrose laurate resulted in the resolution of the eight possible regioisomers (see Fig. 1) with regioisomer assignment based on structural analysis by NMR spectroscopy [8]. R_s-values ranged from 1.25 to 7.07 for sucrose caprate (see Fig. 1a) and from 1.34 to 7.68 for sucrose laurate (see Fig. 1b).

Increasing the eluent acetonitrile concentration decreased the retention times for all the sucrose alkanoate regioisomers (see Fig. 2). However, higher acetonitrile concentrations were required to achieve retention times of sucrose laurate regioisomers similar to those of sucrose caprate. Three regioisomers, 6, 3 and 1′, showed similar elution properties for each of the sucrose alkanoates. The regioisomer 3-*O*-caproyl sucrose was resolved in the concentration range 30–38%, while 3-*O*-lauroyl sucrose was resolved in the

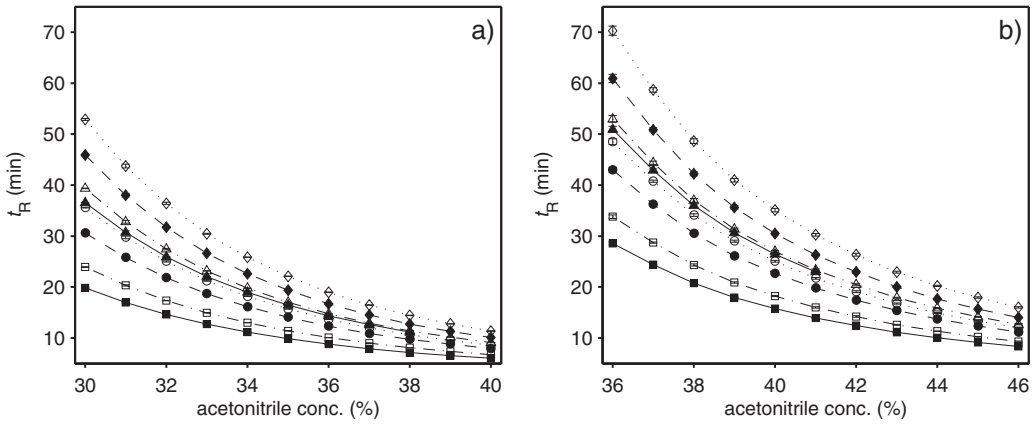


Fig. 2. Regioisomer retention times as function of eluent acetontirile concentrations for isocratic elutions of sucrose caprate (a) and sucrose laurate (b). The regioisomers are arrayed from bottom to top according to elution order: –■ 3′, –·□ 2, –● 4, ··○ 6, –▲ 3, –·△ 1′, –◆ 6′, ··◇ 4′. Error bars indicate standard deviations.

range 36–41%, and above the indicated ranges these regiosiomers co-eluted with their respectively succeeding 1′-regioisomer.

Similar, curved relationships between eluent acetonitrile concentration and retention time, approximated well by cubic polynomials ($R^2 > 0.99$), were observed for all regioisomers. Kendall rank correlation coefficients for second-order retention time differences confirmed the high interdependence of the curves (sucrose caprate regioisomers $\tau \geq 0.81$, sucrose laurate regioisomers $\tau \geq 0.89$, all regioisomers $\tau \geq 0.87$).

3.2. Isocratic elutions with varying flow

Increasing the acetonitrile concentration and the flow rate resulted in increased RD_m and decreased $\tilde{t}'_R$ (see Figs. 3 and 4). The experimental variation of the responses was %RSD = 0.061% for $\tilde{t}'_R$, and %RSD = 2.5% for RD_m.

The regression modelling and statistical analysis showed that although both variables were highly significant for both responses, a change in acetonitrile concentration (I) had a larger effect than a change in flow rate (F) over the ranges studied. The effect of variable I (% ACN) on RD_m was approximately 2.5 times the effect of variable F (mL/min). However, where the effect of F was relatively linear, the effect of I followed a complex relationship, as characterised by statistically significant quadratic and cubic terms (see Fig. 3). The regression model showed %RMSE = 2.9%, which was of the same size-order as the experimental variation. The effect of variable I on $\tilde{t}'_R$ was approximately 1.6 times the effect of variable F, and, while both variables showed a curved relationship with the response, the interaction between the two variables was not statistically significant (see Fig. 4). The regression model showed %RMSE = 0.38%, which was somewhat higher than the experimental variation.

3.3. Step-down gradient elution profiles

In a number of analyses of sucrose caprate the 3′- and 2-regioisomers eluted during the step-down gradient, introducing an additional factor of experimental variation for these regioisomers and resulting in %RSD = 24% for RD_m, however, this did not affect $\tilde{t}'_R$ as much, resulting in %RSD = 2.4%. The experimental design for sucrose laurate avoided elution during the step-down gradient by constraining the maximum summed duration of variables B and D, resulting in lower experimental variation for both RD_m and $\tilde{t}'_R$, which showed %RSD = 2.2% and 1.2%, respectively.

As a minimum, all the regression models explained more than 88% of the variation in the experiments, but most models explained more than 96% of the variation observed (see Table 2). The modelling errors were also of the same size-order as or lower than the experimental variations.

The regression models for average adjusted retention times of the sucrose alkanoates were very similar in the number of terms,

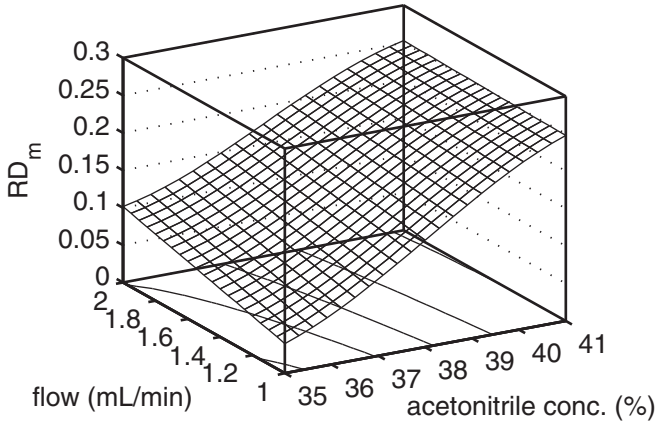


Fig. 3. Response surface and contour plot for regression model of sucrose laurate general resolution deviation (RD_m) as function of eluent acetonitrile concentration and flow rate: $RD_m = 37.22 - 3.027I + 0.08091I^2 - 0.0007090I^3 + 0.3914F - 0.09414IF$; $R^2 = 0.99$, $R^2_{adj.} = 0.99$; %RMSE = 2.9%. The range for RD_m was from 0 to 1, with 0 representing optimal resolution.

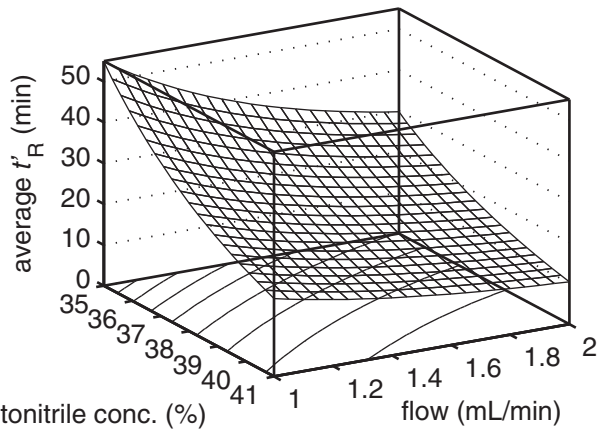


Fig. 4. Response surface and contour plot for regression model of sucrose laurate average adjusted retention time (t'_R) as function of eluent acetonitrile concentration and flow rate: $\log \tilde{t}'_R = 15.84 - 0.4193I + 0.003205I^2 - 1.316F + 0.2350F^2$; $R^2 = 0.98$, $R^2_{adj.} = 0.98$; %RMSE = 0.38%.

Table 2
Regression model evaluation: Coefficients of determination (R^2 and $R^2_{adj.}$) and percent relative root-mean-square error (%RMSE) for models of average adjusted retention time ($\bar{t}'_R$) and general resolution deviation (RD_m) for sucrose alkanoates.

Response	Sucrose caprate		Sucrose laurate	
	$\bar{t}'_R$	RD_m[a]	$\bar{t}'_R$	RD_m[b]
R^2	0.98	0.88	0.99	0.96
$R^2_{adj.}$	0.98	0.84	0.98	0.94
%RMSE	4.1	5.7	1.1	3.2

[a] Data transformed: log RD_m.
[b] Data transformed: $RD_m^{-1.15}$.

but did show some difference in the significance of the terms (see Table 3). The regression model for $\bar{t}'_R$ of sucrose caprate included the additional cubic term for the concentration variable *C* at a low significance level. The linear term for variable *B* was not significant, while all other terms were highly significant. The corresponding model for sucrose laurate included only significant terms, of which all but two were highly significant, the exceptions being the quadratic term for the concentration variable *A* and the interaction between the successive concentration and its duration (*CD*), which both showed a low level of significance.

The regression models for RD_m of the sucrose alkanoates differed in both number of terms and their significance (see Table 3). The regression model for RD_m of sucrose caprate consisted of 14 terms at varying levels of significance. The concentration variables (*A* and *C*) were highly significant as linear and quadratic terms, while the cubic term for *A* showed low significance. The linear terms for duration variables (*B* and *D*) were not significant, but the quadratic term for *D* showed low significance. Of the included two-variable interactions only the interaction between the initial concentration and its duration (*AB*) was significant at medium level, while the three-variable interaction *ACD* showed low significance. The corresponding model for sucrose laurate consisted of 10 terms, also at varying significance levels. The linear terms for all variables were significant, with the initial concentration and its duration (*A* and *B*) at high level and the successive concentration and its duration (*C* and *D*) at low level. The quadratic terms for *A* and *B* were not significant, the cubic terms for these variables showed low significance, while their mutual interaction (*AB*) was highly significant.

The step-up gradient profiles resulted in similar average adjusted retention times and general resolution deviations as the step-down profiles.

3.4. Evaluation of elution efficiency

Efficiency evaluation of sucrose caprate analyses showed that step-down gradient elutions offered combinations of lower RD_m and lower analysis time compared to isocratic elutions (see Fig. 5a). The isocratic elutions in the concentration range 33–35% (analysis time 32, 27 and 24 min, respectively) exhibited the best results. The lowest RD_m was achieved at 33%, with $RD_m = 0.060$, however, elution at 34% reduced the analysis time by 16% with only a 15% increase in RD_m, to 0.069. The gradient profile elutions resulted in a majority of the analyses clustered in the analysis time range from 25 to 31 min, of which twelve elutions showed $RD_m < 0.07$. The lowest values for RD_m were obtained with the acetonitrile concentration variables at 32.5% and 25% (*A* and *C*, respectively) resulting in RD_m in the range 0.052–0.037 and analysis time in the range 29–31 min.

Efficiency evaluation for analyses of sucrose laurate showed that isocratic elutions with increased flow rates offered the most advantageous combinations of RD_m and analysis time (see Fig. 5b). The step-down and the step-up gradient elutions exhibited generally low efficiency with values of RD_m above 0.14 and analysis times above 43 min. The isocratic elutions exhibited a strong negative linear correlation between RD_m and analysis time ($r = -0.99$) with the elutions at acetonitrile concentrations of 36% ($RD_m = 0.052$, 72 min), 37% ($RD_m = 0.10$, 60 min) and 38% ($RD_m = 0.14$, 50 min) offering the most efficient results. For isocratic elutions with varying flow the best efficiency was obtained with flow rates in the range 1.33–2.0 mL/min and concentrations of 35 or 37% acetonitrile, which resulted in analysis time from 33 to 67 min and $RD_m < 0.15$. For these elutions a strong negative linear correlation ($r = -0.99$) between RD_m and analysis time was observed.

4. Discussion

The isocratic elutions of sucrose alkanoates showed similar relationships between the eluent acetonitrile concentration and retention time for all regioisomers of sucrose caprate and sucrose laurate, as confirmed by evaluation of the curvatures using

Table 3
Coefficients and significance of terms for regression models for average adjusted retention time ($\bar{t}'_R$) and general resolution deviation (RD_m) for sucrose alkanoates.

Response	Sucrose caprate		Sucrose laurate	
	$\bar{t}'_R$	RD_m[a]	$\bar{t}'_R$	RD_m[a]
Const.	−48.19	143.8	−78.22	738.3
A	4.078^{***}	-10.76^{***}	4.643^{***}	-52.33^{***}
B	6.370 (*ns*)	−1.328 (*ns*)	10.76^{***}	2.726^{***}
C	-0.8306^{***}	-0.7061^{***}	1.351^{***}	$-3.310 \times 10^{-2*}$
D	2.407^{***}	−5.369 (*ns*)	2.411^{***}	0.1121^{*}
A^2	$-6.148 \times 10^{-2***}$	0.2772^{***}	$-5.927 \times 10^{-2*}$	1.256 (*ns*)
A^3		$-2.432 \times 10^{-3*}$		$-1.007 \times 10^{-2*}$
B^2				0.2714 (*ns*)
B^3				$2.603 \times 10^{-2*}$
C^2	$4.418 \times 10^{-2***}$	$5.643 \times 10^{-3***}$	$-2.338 \times 10^{-2***}$	
C^3	$-7.069 \times 10^{-4*}$			
D^2		$4.794 \times 10^{-2*}$		
AB	-0.1849^{***}	$3.773 \times 10^{-2**}$	-0.2784^{*}	$-9.280 \times 10^{-2***}$
AC		1.113×10^{-2} (*ns*)		
AD		0.1437 (*ns*)		
CD	$-6.590 \times 10^{-2***}$	0.1721 (*ns*)	$-5.527 \times 10^{-2*}$	
ACD		$-4.900 \times 10^{-3*}$		

Significance levels: *ns*, $p > 0.1$.
[a] Data transformed. See Table 2 for details.
* $p \leq 0.1$.
** $p < 0.01$.
*** $p < 0.001$.

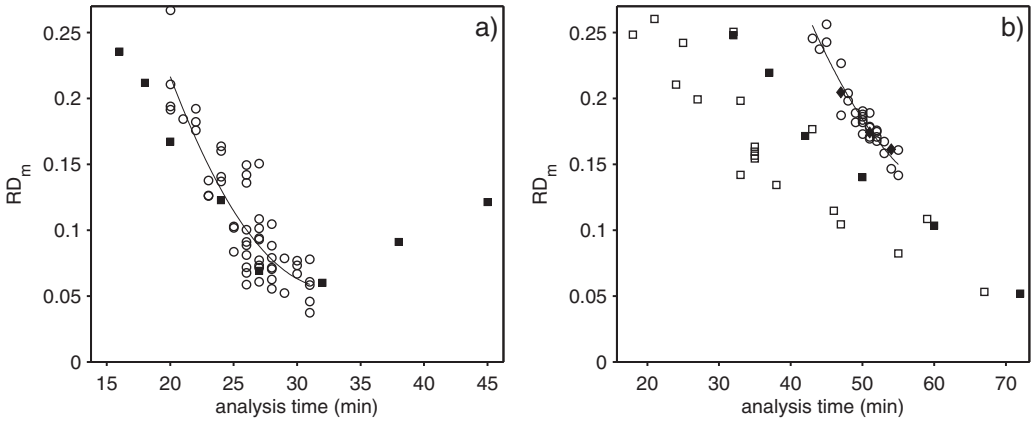


Fig. 5. General resolution deviation (RD_m) and analysis time for isocratic elutions (■), elutions with flow rate variation under isocratic conditions (□), step-down gradient profile elutions (○), and step-up gradient profile elutions (♦) of sucrose caprate (a) and sucrose laurate (b). Lines show the quadratic polynomial fits of the step-down gradient profile elutions ((a) $R^2 = 0.79$, (b) $R^2 = 0.89$).

the approximate second derivatives and Kendall rank correlation coefficients. However, at specific acetonitrile concentrations, increasing retention times were observed for the corresponding regioisomers of sucrose caprate and sucrose laurate, reflecting the difference in the fatty acid chain length between the sucrose alkanoates [12]. Conversely, higher acetonitrile concentrations were required to achieve similar retention times for the corresponding regioisomers of sucrose laurate and sucrose caprate, and furthermore, 3-*O*-lauroyl sucrose was resolved from the succeeding 1′-*O*-lauroyl sucrose at a correspondingly lower acetonitrile concentration than that required for resolving the same pair of sucrose caprate regioisomers.

The results of the isocratic flow-variation elutions conformed to chromatographic theory in showing that lower flow rates give better separation, as caused by increased column efficiency [11]. However, as the difference in effect between flow and acetonitrile concentration was larger for RD_m than for average adjusted retention time, the results indicated that any practical optimisation of separation and length of analysis time should be based on increased flow rates.

Regression modelling of the step-down gradient profiles for the sucrose alkanoates showed that the eluent acetonitrile concentrations were the overall most significant variables for retention time and separation. The models for average adjusted retention time differed in the significance of four terms. Although there were no variables clearly dominating in significance in these models, there was a distinction in the effects between the concentration variables (*A* and *C*) and the duration variables (*B* and *D*), where the former were included as linear terms, exponential terms and part of interaction terms, while the latter variables only contributed as linear or interaction effects. The models were also characterised by simple interactions, in the sense that the only significant interactions were between each concentration variable and the connected duration variable. The models for RD_m showed larger differences between the sucrose alkanoates, in the number of terms and their significance. The model for sucrose caprate was characterised by the concentration variables (*A* and *C*) being the most significant, as the duration variable *B* was significant only as part of its interaction with the initial concentration (*AB*), and the duration variable *D* was significant as the quadratic term and as part of the interaction with both concentration variables (*ACD*). The concentration variables were mostly significant at higher levels as linear and exponential terms and as part of interaction terms. The complexity of this regression model was interpreted as arising from the issue of the experimental variation exhibited by the sucrose caprate RD_m (%RSD = 24%), as also reflected in the lower coefficients of determination for this model compared to the other regression models. Necessarily, the modelling complexity was increased by the retention times for 3′-*O*- and 2-*O*-caproyl sucrose lying within the duration of the step-down gradient, that is, satisfying the relation $t_R < (B + D + 1 + t_0)$, in some, but not all, of the analyses. The regression model for RD_m of sucrose laurate was characterised by the initial concentration and its duration (*A* and *B*) being the most significant variables. These variables were significant as linear and cubic terms and as their mutual interaction term, while the variables *C* and *D* were significant at low level only as linear terms.

The efficiency of the investigated elution strategies was evaluated in terms of RD_m and analysis time, showing that the best results were offered by step-down gradient elution for sucrose caprate and isocratic elution with increased flow for sucrose laurate [13]. The elution efficiency of isocratic elution of sucrose caprate was satisfactory; however, step-down gradient elution offered improved separation at similar analysis time. The most efficient elutions achieved with the acetonitrile concentration variables at 32.5% and 25% (*A* and *C*, respectively) resulted in RD_m decreased by 13–38% and analysis time reduced 3–9% relative to isocratic elution at 33%. These highly similar results for several value combinations of the duration variables were manifestations of the lower significance of these variables, as shown by the regression analysis. The step-down gradient elution strategy that offered improved efficiency for sucrose caprate saw a diminishing return when transferred to sucrose laurate, even to the point that isocratic conditions offered better efficiency. However, isocratic elution with increased flow offered the best efficiency, with elutions at 35 and 37% acetonitrile with flow 1.33–2.0 mL/min resulting in the best combinations of separation and analysis time. In particular, the improvements over isocratic conditions with standard flow (1.0 mL/min) was seen by comparison with the results for elution at 38% acetonitrile, where elution at 37% with flow 2.0 mL/min resulted in reduction of analysis time by 34% and equal RD_m, while elution at 35% with flow 2.0 mL/min reduced RD_m by 29% and analysis time by 6%.

The term for evaluation of peak separation, RD_m, was developed for general applicability by being sample independent in terms of both the number of analytes and their relative amounts. Flexibility of sample composition was provided through the peak size normalisation and allowing for evaluation of any number of peaks. In addition, by normalising peaks to the size of the largest peak, the function avoided overestimation of the resolutions. The previously defined aggregate objective function with a similar purpose [8] was only suitable for comparisons of elution strategies for a specific sample.

5. Conclusion

Efficiency evaluation of elution strategies, in terms of RD_m and analysis time, showed that the best results were offered by step-down gradient elution for sucrose caprate and isocratic elution with increased flow for sucrose laurate. Thus, it was shown that the efficiency of an elution strategy could not be generalised for sucrose alkanoates of increasing alkanoyl chain lengths, as the step-down gradient elution strategy that offered improved results for sucrose caprate saw a diminishing return in efficiency when transferred to sucrose laurate, to the point that the strategy was distinctly outperformed in this context. Step-down gradient elution offered improvements in separation by 13–38% and reductions in analysis time by 3–9% compared to isocratic elution for sucrose caprate, while for sucrose laurate isocratic elution with increased flow showed improvements in separation of up to 29% and reductions of 6–34% in analysis time compared to isocratic elution.

Isocratic elutions of sucrose alkanoates showed similar relationships between eluent acetonitrile concentration and retention time for all regioisomers of sucrose caprate and sucrose laurate, as confirmed by evaluation of the curvatures using approximate second derivatives and Kendall rank correlation coefficients.

Regression modelling and statistical analysis showed that acetonitrile concentration and flow rate were highly significant for both average adjusted retention time and RD_m for sucrose laurate. For both responses the effect of changes in acetonitrile concentration was larger than the effect of changes in flow rate, over the ranges studied.

Regression modelling of the step-down gradient profiles for the sucrose alkanoates showed that the eluent acetonitrile concentrations were the overall most significant variables for retention time and separation. The models for average adjusted retention time of sucrose caprate and sucrose laurate showed only a few differences in the significance levels of terms, while the models for RD_m showed larger differences between the sucrose alkanoates, in both the number of terms and their significance.

General resolution deviation for multiple peaks (RD_m) was developed for sample-independent evaluation of overall separation of any number of peaks in chromatographic analysis.

Acknowledgement

The present research was funded by the Danish Council for Independent Research.

References

[1] M. Ferrer, J. Soliveri, F.J. Plou, N. Lopez-Cortes, D. Reyes-Duarte, M. Christensen, J.L. Copa-Patino, A. Ballesteros, Enzyme Microb. Technol. 36 (2005) 391.
[2] G.J. Puterka, W. Farone, T. Palmer, A. Barrington, J. Econ. Entomol. 96 (2003) 636.
[3] N. Garti, V. Clement, M. Leser, A. Aserin, M. Fanun, J. Mol. Liq. 80 (1999) 253.
[4] E.N. Vulfson, Trends Food Sci. Technol. 4 (1993) 209.
[5] F.J. Plou, M.A. Cruces, M. Ferrer, G. Fuentes, E. Pastor, M. Bernabe, M. Christensen, F. Comelles, J.L. Parra, A. Ballesteros, J. Biotechnol. 96 (2002) 55.
[6] I. Perez-Victoria, A. Zafra, J.C. Morales, Carbohydr. Res. 342 (2007) 236.
[7] S. Ritthitham, R. Wimmer, A. Stensballe, L.H. Pedersen, J. Chromatogr. A 1216 (2009) 4963.
[8] A. Lie, R. Wimmer, L.H. Pedersen, J. Chromatogr. A 1281 (2013) 67.
[9] S.S. Shapiro, M.B. Wilk, Biometrika 52 (1965) 591.
[10] D.C. Montgomery, Design and Analysis of Experiments, John Wiley & Sons, Singapore, 2012, pp. 730.
[11] R. Tijssen, in: E. Katz, R. Eksteen, P. Schoenmakers (Eds.), The Mechanisms and Importance of Zone-Spreading. Handbook of HPLC, Marcel Dekker, New York, NY, USA, 1998, p. 55.
[12] M. Lisa, F. Lynen, M. Holcapek, P. Sandra, J. Chromatogr. A 1176 (2007) 135.
[13] S. Kromidas, in: S. Kromidas (Ed.), Fundamentals of Optimization: Principles of the Optimization of HPLC Illustrated by RP-Chromatography. HPLC Made to Measure: A Practical Handbook for Optimization, Wiley-VCH Verlag GmbH & Co. KGaA, 2006, p. 3.

3 APPEARANCE AND DISTRIBUTION OF REGIOISOMERS IN METALLO- AND SERINE-PROTEASE-CATALYSED ACYLATION OF SUCROSE IN *N*,*N*-DIMETHYLFORMAMIDE

Lie A., Meyer A.S., and Pedersen L.H.

Journal of Molecular Catalysis B: Enzymatic (2014), 106, 26-31

Journal of Molecular Catalysis B: Enzymatic 106 (2014) 26–31

Contents lists available at ScienceDirect

Journal of Molecular Catalysis B: Enzymatic

journal homepage: www.elsevier.com/locate/molcatb

Appearance and distribution of regioisomers in metallo- and serine-protease-catalysed acylation of sucrose in *N*,*N*-dimethylformamide

Aleksander Lie[a], Anne S. Meyer[b], Lars Haastrup Pedersen[a,*]

[a] Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
[b] Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søltofts Plads, Building 227, 2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Article history:
Received 3 February 2014
Received in revised form 11 April 2014
Accepted 23 April 2014
Available online 2 May 2014

Keywords:
Sucrose fatty acid ester
Regioisomeric distribution
Protease-catalysed acylation
Hydrophilic solvent
Non-enzymatic catalysis

ABSTRACT

The appearance and distribution of monoester regioisomers were investigated in the virtually irreversible acylation of sucrose with the enol ester, vinyl laurate, as acyl donor catalysed by serine proteases and a metalloprotease in the hydrophilic, aprotic solvent *N*,*N*-dimethylformamide. Sucrose laurate was obtained in yields from 12 to 53% after 48 h under different catalytic conditions. The serine protease ALP-901, derived from a *Streptomyces* sp., produced the highest yield at this reaction time, while reaction with the zinc-protease thermolysin achieved the overall highest yield (63%) after 6 h, with only monoesters synthesised. The total conversion of sucrose after 48 h ranged from 19 to 96%. The highest degree of conversion was observed in the reaction with thermolysin, while the reactions without protein and with ALP-901 resulted in 82% and 66% sucrose conversion, respectively. 2-*O*-Lauroyl sucrose was the most abundant monoester regioisomer synthesised and the highest concentration observed was 23.7 mM after 24 h in the thermolysin-catalysed reaction. The highest concentration of 2-*O*-lauroyl sucrose detected in the reaction catalysed by ALP-901 was 19.0 mM, while it was 17.0 mM the reaction without protein, both after 48 h. The detected appearance of the sucrose laurate regioisomers largely corresponded to the apparent rates of formation, and 2-*O*-lauroyl sucrose was among the first regioisomers to appear in all reactions. The observed sucrose laurate regioisomeric distribution after 48 h (2:3:4:6:1′:3′) was 72:5:2:1:7:14 in the reaction catalysed by ALP-901, and 74:5:2:1:7:13 in the reaction without protein. In the reaction catalysed by thermolysin the distribution was 71:5:2:–:9:13 after 6 h and 86:8:–:–:4:3 after 48 h of reaction. The esterification of sucrose with vinyl laurate without protein in the reaction mixture appeared to be catalysed in the presence of aluminosilicate molecular sieves. Non-catalytic protein in the reaction medium seemed to lower the catalytic activity of the molecular sieves.

1. Introduction

Sucrose fatty acid esters are widely used in the food, cosmetics, pharmaceutical and detergent industries, as they can be produced with a wide range of emulsifying, dispersing and solubilising properties [1,2]. These non-ionic surfactants are non-toxic and biodegradable, and some of them have shown antimicrobial effects [1,3,4]. Currently sucrose alkanoates are synthesised by conventional chemical processes at high temperatures, resulting in low regiospecificity and side reactions leading to alterations of product properties [2,5,6]. Biocatalysis has been shown as a strong alternative to conventional chemical synthesis, offering improved activity and selectivity while not needing the protection and deprotection steps often required in chiral and regioselective organic synthesis. In particular, lipases and proteases have been applied to the synthesis of sucrose alkanoates in organic solvents [7]. Lipases have been shown to catalyse esterification by reversed hydrolysis reactions and exhibit selectivity towards the primary hydroxyl groups of sucrose (6, 1′ and 6′) [8,9], while proteases catalyse transesterification reactions and exhibit selectivity towards primary or secondary hydroxyl groups in sucrose [10,11]. Hydrophilic, aprotic solvents, including *N*,*N*-dimethylformamide (DMF) and dimethylsulfoxide (DMSO) are excellent solvents for carbohydrates but act as protein denaturants affecting enzyme activity and stability. Proteases have shown improved activity and selectivity with

* Corresponding author. Tel.: +45 9940 8497; fax: +45 9814 1808.
E-mail addresses: alie@bio.aau.dk (A. Lie), am@kt.dtu.dk (A.S. Meyer), lhp@bio.aau.dk (L.H. Pedersen).

http://dx.doi.org/10.1016/j.molcatb.2014.04.017
1381-1177/

DMSO as co-solvent [12,13], while lipases have shown sensitivity to hydrophilic solvents, losing activity and exhibiting complete inactivation at DMSO concentrations of 20–30% and above [14,15]. Subtilisin has shown selectivity towards 1′-OH and 6-OH of sucrose in anhydrous DMF as well as in mixtures of buffer and DMSO [16–18], while the metalloprotease thermolysin catalysed the acylation of sucrose primarily at the 2-OH [19] and the alkaline protease AL-89 catalysed the esterification of 2-OH and promoted acyl migration to the 3-regioisomer of sucrose alkanoates [13]. The recent developments in methods for quantitative analysis of sucrose laurate using RP-HPLC with charged aerosol detection (CAD) [20] made more accurate studies of the esterification processes possible, and in particular offered increased sensitivity in the analysis of the distribution of regioisomers. In this work, we have investigated catalytic activity and appearance and distribution of monoester regioisomers in the irreversible acylation of sucrose with the enol ester vinyl laurate as acyl donor catalysed by serine proteases and a metalloprotease in the hydrophilic, aprotic solvent DMF.

2. Materials and methods

2.1. Materials

All solvents were of HPLC grade, unless otherwise indicated. Methanol (MeOH), trifluoroacetic acid (TFA, reagent grade) and vinyl laurate (GC grade) were from Sigma–Aldrich, acetonitrile (ACN) was from Panreac Quimica, while *N*,*N*-dimethylformamide (DMF, GC grade) was from Fluka. Water (H_2O) of Type I purity (ELGA PURELAB flex) was degassed by sonication for 20 min (Bransonic 2510E-MT, water bath). Sucrose was from Nordic Sugar A/S, while molecular sieves (0.3 nm, aluminosilicate beads ~2 nm) were from Merck. Enzyme and protein formulations used were ALP-901 (crude formulation of ALP-101 from *Streptomyces* sp., Toyobo), subtilisin A (*Bacillus licheniformis*, lyophilised, Novozymes A/S), Alcalase 3.0 T (granulate of subtilisin A, Novozymes A/S), thermolysin (*Bacillus thermoproteolyticus*, lyophilised, Fluka) and casein (technical grade, Sigma–Aldrich).

2.2. Sucrose laurate synthesis reactions

Sucrose was powdered with mortar and pestle and dried at 120 °C for 24 h, while DMF and vinyl laurate were dried over molecular sieves (40 g/L) for at least 24 h. Reactions were performed by dissolving sucrose (50 mM) in DMF (5 mL) containing molecular sieves (40 g/L) in serum flasks (25 mL) with stirring magnets for ca. 10 min in a stirring block thermostat (50 °C, 600 rpm, VARIOMAG 15.5 with TELEMODUL 40 CT). Enzyme or protein formulation (10 g/L) was added where appropriate, vinyl laurate (400 mM) was added and the flasks stoppered with rubber caps and sealed with parafilm. Reactions with ALP-901, alcalase, and subtilisin were run in three parallels, and the reaction with thermolysin was run in two parallels. Control reactions were performed with non-catalytic protein (casein) and without protein; the reaction with casein was run in three parallels, as was the reaction without protein to 6 h, while a single reaction without protein was run to 48 h. The chemical background reaction was accounted for by a reaction with neither protein nor molecular sieves, performed in duplicate.

After 2, 4, 6, 24 and 48 h reaction time, samples (100 μL) were withdrawn from the reaction mixture using a syringe (1 mL) and transferred to centrifuge tubes (1.2 mL). After centrifugation (14,000 rpm, 3+ min) supernatant aliquots (50 μL) were diluted with methanol (ratio 1:9, 1:2 for the reaction without protein) for analysis by RP-HPLC.

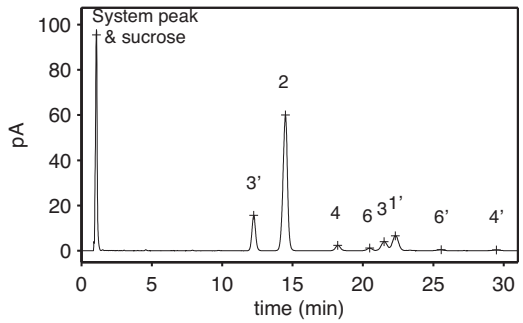


Fig. 1. Chromatogram of reaction mixture from reaction catalysed by thermolysin after 6 h. Peaks annotated with regioisomer substitution position. Retention times: sucrose, 1.1; sucrose laurate regioisomers, 12.2, 14.5, 18.2, 20.5, 21.5, 22.3, 25.6, 29.5.

2.3. RP-HPLC analysis

Samples (10 μL injections) were analysed using a Dionex UltiMate 3000 system (Thermo Scientific) with Corona Veo CAD charged aerosol detector (Thermo Scientific, N_2 at 61.5 psi, nebuliser at 50 °C). Stationary phase was a Symmetry C_{18} column (250 mm × 4.6 mm, 5 μm particle size, Waters) maintained at 45 °C, and eluents consisted of mixtures of water and acetonitrile, both with 0.05% (v/v) TFA, with flow rate 2.0 mL/min. The chromatographic data were collected using Chromeleon software (version 7.2, Thermo Scientific).

Sucrose laurate regioisomer analysis was conducted using isocratic elution with 37% acetonitrile in water for 31 min, as previously optimised for these analytes [20] (see Fig. 1). Oligoester analysis was conducted by elution with 50% ACN held to 5 min, followed by a gradient to 100% ACN at 10 min, and 100% ACN held to 15 min.

Standard curves for sucrose and 6-*O*-lauroyl sucrose were constructed from the analysis of sample series over the respective ranges of 0.05–60 and 0.06–90 μg injected analyte. CAD exhibits characteristic non-linear mass–response relations expressed by the equation:

$$A = aM^b \tag{1}$$

where A is the area response of the detector, M is the injected mass of analyte, and a and b are values specific to analytes and chromatographic conditions. Linear regression was applied to logarithmically transformed peak areas, and from the linearised mass–response relation:

$$\log A = b \log M + \log a \tag{2}$$

the coefficient b and the term $\log a$ were determined for each analyte [21]. By rearranging Eq. (1) and inserting the determined values, the injected mass of analyte (M, ng) were determined from peak area (A, pA s) according to the equations:

$$M \pm 2.1\% = \frac{A^{1.89}}{1.565} \tag{3}$$

for sucrose, and

$$M \pm 3.5\% = 4.20A^{1.09} \tag{4}$$

for sucrose laurate.

2.4. Mass spectrometry

Sucrose laurate product (lyophilised, 3.3 mg) in methanol (100 μL) was analysed by matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry, as

Table 1
Initial rates of formation for 2-*O*-lauroyl sucrose.

Enzyme formulations (w/mol sieves)				Controls (w/mol sieves)		Control (w/o mol sieves)
Thermolysin	ALP-901	Alcalase	Subtilisin	Casein	No protein	Background
108 ± 6	22.2 ± 1.9	1.78 ± 0.17	0.765 ± 0.234	0.981 ± 0.020	3.24 ± 0.63	0.238 ± 0.019

Initial rates (μM/min) were calculated from 2-h samples, except for the casein reaction where it was calculated from the 4-h sample.

previously described [19], using DHB (2,5-dihydroxybenzoic acid) as matrix.

3. Results and discussion

All protease-catalysed reactions, as well as the control reactions with casein and without protein, resulted in the synthesis of sucrose laurate. Mass spectrometry of the sucrose laurate product gave characteristic signals from the molecular ion adducts with Na and K (*m*/*z* 547 and 563, respectively), corresponding to the observations of Wang et al. [22] for glucose, fructose and sucrose. 2-*O*-Lauroyl sucrose was the most abundant monoester synthesised, and its concentration increased over the first 6 h of reaction time in all reactions (see Fig. 2). The initial rates of formation for this regioisomer ranged from 0.238 to 108 μM/min (see Table 1), with the control reaction with neither protein nor molecular sieves showing the lowest rate. This reaction was considered as the chemical background reaction. The three conditions exhibiting highest initial rates were thermolysin, ALP-901 and no protein, and the rates with thermolysin and ALP-901 were, respectively, more than 450 and 90 times the rate of the chemical background reaction, and about 33 and 7 times the rate of the reaction without protein (but with molecular sieves).

Compared to the remaining reactions where molecular sieves were included, the initial rate of the reaction without protein was about 2 times that with alcalase, about 3.5 times the rate with casein and about 4 times the rate with subtilisin, indicating the presence of catalytic functionality other than enzyme under these reaction conditions. Ritthitham et al. [13] showed Celite to have catalytic activity in the reaction between sucrose and vinyl stearate in DMF/DMSO (1:1, v/v). The initial rate of formation for 2-*O*-stearoyl sucrose was reported to be 0.17 μM/min in medium without Celite, 1.7 μM/min with 10 g/L Celite in the medium, while it increased to 11.7 μM/min with 100 g/L. The molecular sieves employed in the present investigation were made of sodium aluminosilicate, a material similar to Celite, and the initial rate observed in the reaction without protein and the concentration of molecular sieves (40 g/L) corresponded well with their observations. In conjunction with these results, the lower initial rate with the non-catalytic protein casein added to the reaction, seemed to be due to the presence of protein and may be caused by an interaction between protein and the molecular sieves reducing the catalytic effect.

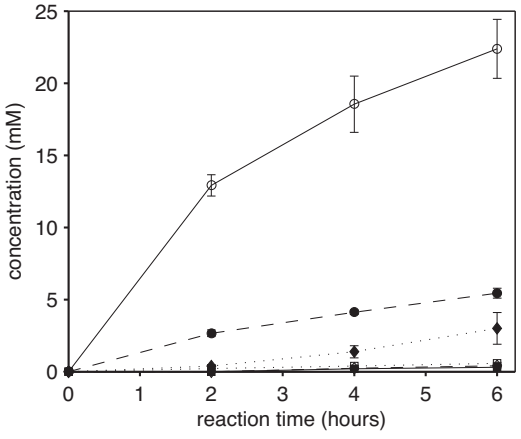


Fig. 2. Progress curves for 2-*O*-lauroyl sucrose. Reaction conditions: (–○) thermolysin, (–●) ALP-901, (···□) alcalase, (–■) subtilisin, (–◊) casein, (···♦) no protein. Error bars indicate standard deviation, determined from triplicate reactions, except for thermolysin (duplicates).

The lower initial rates of the subtilisin formulations compared to the reaction without protein, indicated decreased catalytic activity in these reactions, similar to what was observed in the presence of the non-catalytic protein, casein. Thus, catalytic activity specifically ascribable to subtilisin in either of the two formulations could not be distinguished from that of the non-catalytic protein, casein. Substrate inhibition, particularly with vinyl laurate, could decrease the activity of the enzymes, although subtilisin has been reported to catalyse the acyl-transfer between vinyl laurate and sucrose in DMF–pyridine (1:1, v/v) at higher acyl donor concentrations than employed in the present investigation [10]. Subtilisin and other alkaline proteases have also been reported to exhibit high stability in pure DMF, in fact, higher than in other commonly used hydrophilic, aprotic solvents, such as dimethyl sulfoxide (DMSO), and in some cases higher than in water [13,23], although the granulate of subtilisin A (Alcalase) was observed to dissolve in the reaction medium. In addition, the tautomerisation by-product of acyl transfer with vinyl esters, acetaldehyde, can act as an alkylating agent through formation of Schiff's bases with lysine residues in the protein, which destabilises some enzymes [24], although the presence of molecular sieves has been indicated as beneficial through trapping of the acetaldehyde [25].

The concentration of sucrose was observed to decrease progressively during the course of all the reactions, and the total conversion of sucrose after 48 h ranged from 19 to 96% (see Table 2). The highest degree of conversion was observed in the reaction with thermolysin, while the reactions with no protein and ALP-901 showed degrees of conversion above 50%. The sucrose laurate yield after 48 h ranged from 12 to 53%, and the reaction with ALP-901 resulted in the highest yield. To allow for the complete ester substitution of sucrose, 8 equiv. of acyl donor were used in the reactions. The yield of sucrose laurate compared to the substrate conversion in the reactions with thermolysin, ALP-901 and without protein indicated the formation of di- and oligoesters at the expense of monoesters. The presence of oligoesters was observed after 48 h in these reactions, and the detected amounts of mono- and oligoesters corresponded to the degree of sucrose conversion. Previous reports suggest that a lower ratio of acyl donor to acyl acceptor would increase the selectivity towards monoester product [10,26].

Table 2
Sucrose conversion and sucrose laurate yields after 48 h.

Reaction	Sucrose conversion (%)	Sucrose laurate yield (%)
Thermolysin	96 ± 3	29 ± 2
ALP-901	66 ± 5	53 ± 4
Alcalase	34 ± 5	44 ± 5
Subtilisin	19 ± 8	12 ± 3
Casein	21 ± 9	34 ± 9
No protein	82 ± 2	46 ± 10

Conversion expressed as ratio of concentration decrease to initial concentration, and yield expressed as ratio of detected concentration to theoretical maximum.

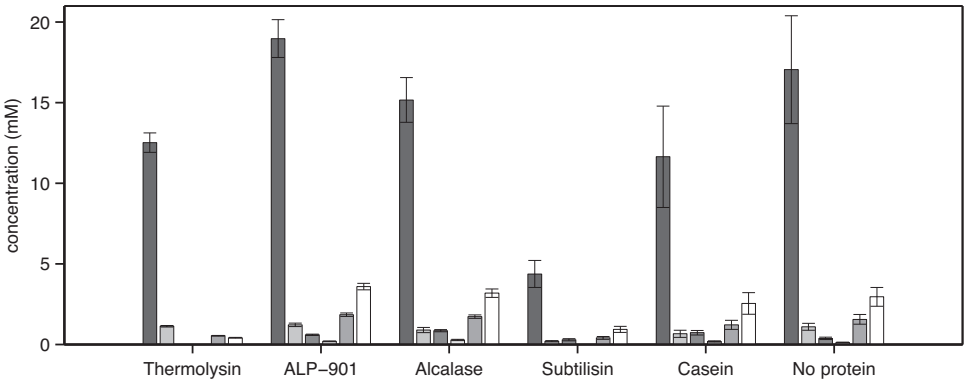


Fig. 3. Catalyst effect on sucrose laurate regioisomer concentration in reaction mixture after 48 h. Regioisomers of sucrose laurate: (■) 2, (■) 3, (■) 4, (■) 6, (■) 1′, □ 3′. The 4′- and 6′-regioisomers were not detected except in trace amounts in the reaction without protein. Error bars indicate standard deviations, determined from triplicate reactions, except for thermolysin (duplicates).

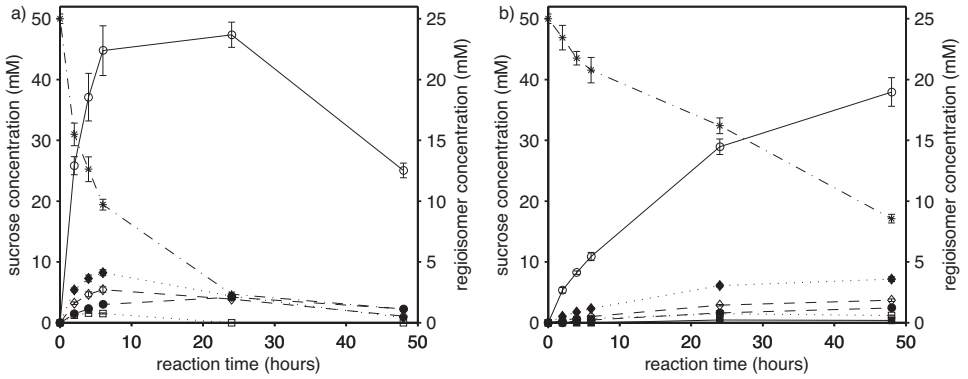


Fig. 4. Progress curves for the reactions catalysed by (a) thermolysin and (b) ALP-901. Sucrose: (–•✱); sucrose laurate regioisomers: (–○) 2, (–●) 3, (···□) 4, (–■) 6, (–◇) 1′, (···◆) 3′. Error bars indicate standard deviations determined from duplicates for thermolysin and triplicates for ALP-901.

The total concentration of sucrose laurate observed after 48 h differed between the reaction conditions, however, 2-*O*-lauroyl sucrose was the most abundant regioisomer in all reactions (see Fig. 3). The regioisomeric distribution of sucrose monoesters after 48 h was similar in the reactions with alcalase, subtilisin and casein, while it differed somewhat in the reactions with ALP-901 and without protein, whose distributions were similar to each other. In the background control, very low concentrations of only the 2- and 3′-regioisomer were detected during 6 h of reaction. The reaction with thermolysin exhibited a unique regioisomeric distribution with lower proportion of the 3′-regioisomer and the 4-regioisomer not detected (see Table 3). The detected appearance of the sucrose laurate regioisomers largely corresponded to the apparent rates of formation, based on regioisomeric distribution and concentration after 48 h.

The reaction progress over 48 h for the two enzyme-catalysed reactions with the highest initial rates was compared (see Fig. 4). In the thermolysin-catalysed reaction the highest concentration of 2-*O*-lauroyl sucrose (23.7 mM) was observed at 24 h (see Fig. 4a). Other regioisomers detected from 2 h, in descending order of concentration, were 3′, 1′, 3 and 4 (see Table 3). The 4-regioisomer concentration peaked at 4 h and could no longer be detected from 24 h. The concentrations of regioisomers 3′ and 1′ peaked at 6 h, while the concentration of the 3-regioisomer was highest at 24 h. The overall highest yield of sucrose laurate (63 ± 5%) was detected in the reaction with thermolysin already after 6 h (see Fig. 4a),

Table 3
Catalyst effect on appearance and distribution of sucrose laurate regioisomers.

Reaction	Reaction time					Regioisomeric distribution after 48 h (2:3:4:6:1′:3′)[a]
	2 h	4 h	6 h	24 h	48 h	
Thermolysin	2, 3, 4, 1′, 3′					86:8:–:–:4:3
ALP-901	2, 4, 1′, 3′	3, 6				72:5:2:1:7:14
Alcalase	2	1′, 3′		3, 4, 6		69:4:4:1:8:14
Subtilisin	2		3′	4, 1′	3	71:3:5:–:7:15
Casein		2, 3′		3, 4, 1′	6	69:4:4:1:7:15
No protein	2, 3′	3, 4, 6, 1′		[b]	4′, 6′	74:5:2:1:7:13

Regioisomer identity is indicated under the reaction time of first detection for each reaction condition.
[a] The regioisomers 4′ and 6′ were only detected in trace amounts in the reaction without protein. Sums of ratios can differ from 100 due to rounding.
[b] The reaction without protein was not sampled after 24 h.

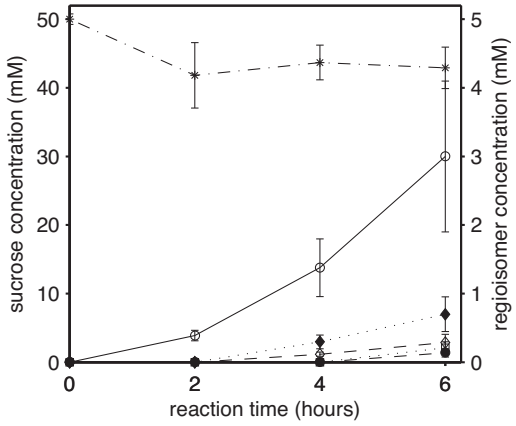


Fig. 5. Progress curves for the reaction without protein. Sucrose: (-·✱); sucrose laurate regioisomers: (–○) 2, (–●) 3, (···□) 4, (–■) 6, (–◇) 1′, (···◆) 3′. Error bars indicate standard deviations determined from triplicate reactions.

where the regioisomeric distribution was 71:5:2:9:13 (2:3:4:1′:3′, 6, 6′ and 4′ not detected) and the degree of sucrose conversion indicated only monoester product was present. The subsequent decline in sucrose laurate concentration was due to oligoester formation, as previously mentioned.

In the reaction catalysed by ALP-901 the concentration of 2-*O*-lauroyl sucrose increased progressively during the reaction to 19.0 mM at 48 h (see Fig. 4b). Other regioisomers detected at 2 h, in descending order of concentration, were 3′, 1′ and 4, while the 3- and 6-regioisomers were detected from 4 h (see Table 3). The concentrations of regioisomers 4 and 6 peaked at 24 h, while the other regioisomers detected were observed in highest concentration after 48 h of reaction.

The reaction progress over 6 h for the control reaction without protein was also considered (see Fig. 5). The 2-regioisomer was the most abundant at 3.00 mM after 6 h, and the regioisomeric distribution at this time, 69:3:5:1:7:16 (2:3:4:6:1′:3′), was very similar to the distributions observed in the reactions with alcalase, subtilisin and casein after 48 h (see Table 3). The 3′-regioisomer was detected from 2 h and the regioisomers 1′, 4, 6 and 3, in descending order of concentration, were detected from 4 h. The concentration of sucrose varied irregularly between the time points and the standard deviations were also large, possibly as a result of the samples analysed containing sucrose concentrations close to the upper limit of the detector range. After 48 h, all regioisomers were detected in the sample and the highest concentrations of all regioisomers were observed, with 17.0 mM of 2-*O*-lauroyl sucrose detected.

The observed selectivity of both thermolysin and ALP-901 was predominantly towards formation of the 2-regioisomer. This corresponded to previous reports on thermolysin-catalysed acyl transfer reactions in hydrophilic, aprotic solvents [19,27], and reports on catalysis by another alkaline protease, AL-89, in vinyl fatty acid acyl transfers to sucrose in DMF-DMSO (1:1, v/v) [13]. Previously reported subtilisin-catalysed reactions with short-chained acyl donors in various DMF–co-solvent mixtures, have shown selectivity for acylation at the 1′-position of sucrose [11,16,28], however, as outlined above, the catalytic activity of the subtilisin preparations employed in the present investigation appeared to be negligible and the regioisomeric distribution. At 48 h the regioisomeric distribution of the thermolysin catalysed reaction differed significantly from the remaining reactions, while the proportion of the 4-regioisomer of the reactions with ALP-901 and without protein was half that of the three slower reactions. Nuclear magnetic resonance studies of sucrose in DMSO solution have showed intramolecular hydrogen bonds between 2-OH and both 1′-OH and 3′-OH [29], which have conformation-stabilising effects. Computer modelling of molecular electrostatic potential profiles showed the 2-OH substituent to be the most electropositive hydroxyl group, and experimental confirmation by benzylation in DMF resulted in the order of reactivity 2-OH > 1′-OH > 3′-OH [30]. Acid chloride acylation in water have also confirmed the 2-OH as the most acidic and reactive substituent [31,32]. Molinier et al. [33] showed that acyl migration in sucrose monoesters is catalysed by aqueous, hydrophilic organic media, resulting in rapid conversion from the 2- to the 3-regioisomer, which has also been shown to be catalysed by the protease AL-89 [13], and slower further conversion to the 6-regioisomer. Even slower migrations favour the primary hydroxyl groups 6, 1′ and 6′. The resulting regioisomeric distribution of sucrose acylation does not necessarily express the relative reactivity of the hydroxyl groups, but the intricate competition between esterification and migration [32]. In the current reactions, the regioisomeric distribution conformed largely to the reactivity predicted by the electrostatic potential with the major products being the 2-, 3′- and 1′-regioisomers, indicating that acyl migration was not favoured.

The distinctly different product compositions after 48 h in the reaction with thermolysin seemed to be a result of the overall higher reaction rate and the reaction thus progressing further over the reaction time studied. The reaction progress over 48 h for ALP-901 resembled that over 6 h for thermolysin, as supported by the similarity of the sucrose laurate regioisomeric distribution and degree of sucrose conversion at these respective reaction times, and by the decline in the concentration of some regioisomers from after 4 h with thermolysin and after 24 h with ALP-901.

4. Conclusion

The acylation of sucrose with vinyl laurate acyl donor in the hydrophilic, aprotic solvent *N*,*N*-dimethylformamide resulted in the formation of sucrose laurate in yields from 12 to 53% after 48 h under different catalytic conditions. The serine protease ALP-901 produced the highest yield at this reaction time, while the reaction with thermolysin achieved the overall highest yield (63%) after 6 h, with only monoesters synthesised. 2-*O*-Lauroyl sucrose was the most abundant monoester regioisomer synthesised, it was among the first regioisomers to appear and its concentration increased over the first 6 h in all reactions. The highest concentration of 2-*O*-lauroyl sucrose observed was 23.7 mM after 24 h in the thermolysin-catalysed reaction, while in the reaction catalysed by ALP-901 and in the reaction without protein, it was 19.0 mM and 17.0 mM after 48 h, respectively. The initial rates of formation of the reactions catalysed by thermolysin and ALP-901 were more than 450 and 90 times the rate of the chemical background reaction, respectively.

The detected appearance of the sucrose laurate regioisomers largely corresponded to the apparent rates of formation, based on initial rates of formation for the 2-regioisomer and monoester concentration and regioisomeric distribution after 48 h. The observed sucrose laurate regioisomeric distribution after 48 h (2:3:4:6:1′:3′) was 72:5:2:1:7:14 in the reaction catalysed by ALP-901, and 74:5:2:1:7:13 in the reaction without protein. In the reaction catalysed by thermolysin, the distribution was 71:5:2:–:9:13 after 6 h, while it was 86:8:–:–:4:3 after 48 h. The concentration of sucrose laurate was observed to decline with reaction times above 6 h in the reaction with thermolysin due to oligoester formation.

The reaction between sucrose and vinyl laurate with no protein in the reaction mixture appeared to be catalysed by aluminosilicate molecular sieves present in the reaction medium, with an initial rate of formation 13 times higher than that of the chemical background reaction. Non-catalytic protein in the reaction medium seemed to lower the catalytic activity of the molecular sieves.

Acknowledgements

We are grateful to Associate Prof. Allan Stensballe at the Department of Health Science and Technology, Aalborg University, for performing the mass spectrometry. The present research was funded by the Danish Council for Independent Research.

References

[1] M. Ferrer, J. Soliveri, F.J. Plou, N. Lopez-Cortes, D. Reyes-Duarte, M. Christensen, J.L. Copa-Patino, A. Ballesteros, Enzyme Microb. Technol. 36 (2005) 391–398.
[2] F.J. Plou, M.A. Cruces, M. Ferrer, G. Fuentes, E. Pastor, M. Bernabe, M. Christensen, F. Comelles, J.L. Parra, A. Ballesteros, J. Biotechnol. 96 (2002) 55–66.
[3] M. Habulin, S. Sabeder, Z. Knez, J. Supercrit. Fluids 45 (2008) 338–345.
[4] T. Watanabe, S. Katayama, M. Matsubara, Y. Honda, M. Kuwahara, Curr. Microbiol. 41 (2000) 210–213.
[5] E. Castillo, F. Pezzotti, A. Navarro, A. Lopez-Munguia, Biotechnol. J. 102 (2003) 251–259.
[6] J.F. Robyt, Essentials of Carbohydrate Chemistry, Springer, New York, 1998.
[7] Y.G. Shi, J.R. Li, Y.H. Chu, J. Chem. Technol. Biotechnol. 86 (2011) 1457–1468.
[8] G. Carrea, S. Riva, Angew. Chem. Int. Ed. 39 (2000) 2226–2254.
[9] N.R. Pedersen, R. Wimmer, J. Emmersen, P. Degn, L.H. Pedersen, Carbohydr. Res. 337 (2002) 1179–1184.
[10] N.R. Pedersen, R. Wimmer, R. Matthiesen, L.H. Pedersen, A. Gessesse, Tetrahedron: Asymmetry 14 (2003) 667–673.
[11] S. Riva, J. Chopineau, A.P.G. Kieboom, A.M. Klibanov, J. Am. Chem. Soc. 110 (1988) 584–589.
[12] O. Almarsson, A.M. Klibanov, Biotechnol. Bioeng. 49 (1996) 87–92.
[13] S. Ritthitham, R. Wimmer, A. Stensballe, L.H. Pedersen, J. Mol. Catal. B: Enzyme 59 (2009) 266–273.
[14] P. Degn, W. Zimmermann, Biotechnol. Bioeng. 74 (2001) 483–491.
[15] D. Reyes-Duarte, N. Lopez-Cortes, M. Ferrer, F.J. Plou, A. Ballesteros, Biocatal. Biotransform. 23 (2005) 19–27.
[16] P. Potier, A. Bouchu, J. Gagnaire, Y. Queneau, Tetrahedron: Asymmetry 12 (2001) 2409–2419.
[17] S. Riva, M. Nonini, G. Ottolina, B. Danieli, Carbohydr. Res. 314 (1998) 259–266.
[18] T. Watanabe, R. Matsue, Y. Honda, M. Kuwahara, Carbohydr. Res. 275 (1995) 215–220.
[19] N.R. Pedersen, P.J. Halling, L.H. Pedersen, R. Wimmer, R. Matthiesen, O.R. Veltman, FEBS Lett. 519 (2002) 181–184.
[20] A. Lie, L.H. Pedersen, J. Chromatogr. A 1311 (2013) 127–133.
[21] T. Vehovec, A. Obreza, J. Chromatogr. A 1217 (2010) 1549–1556.
[22] C. Wang, J. Li, S. Yao, Y. Guo, X. Xia, Anal. Chim. Acta 604 (2007) 158–164.
[23] H. Ogino, H. Ishikawa, J. Biosci. Bioeng. 91 (2001) 109–116.
[24] H.K. Weber, H. Stecher, K. Faber, Biotechnol. Lett. 17 (1995) 803–808.
[25] K. Faber, Biotransformations in Organic Chemistry: A Textbook, 5th ed., Springer, Heidelberg, 2004.
[26] M. Ferrer, M.A. Cruces, M. Bernabe, A. Ballesteros, F.J. Plou, Biotechnol. Bioeng. 65 (1999) 10–16.
[27] I. Perez-Victoria, J.C. Morales, Tetrahedron 62 (2006) 2361–2369.
[28] J.O. Rich, B.A. Bedell, J.S. Dordick, Biotechnol. Bioeng. 45 (1995) 426–434.
[29] J.C. Christofides, D.B. Davies, J. Chem. Soc.: Chem. Commun. (1985) 1533–1534.
[30] F.W. Lichtenthaler, S. Immel, P. Pokinskyj, Liebigs Ann. 1995 (1995) 1939–1947.
[31] Y. Queneau, S. Chambert, C. Besset, R. Cheaib, Carbohydr. Res. 343 (2008) 1999–2009.
[32] S. Thevenet, A. Wernicke, S. Belniak, G. Descotes, A. Bouchu, Y. Queneau, Carbohydr. Res. 318 (1999) 52–66.
[33] V. Molinier, K. Wisnienwski, A. Bouchu, J. Fitremann, Y. Queneau, J. Carbohydr. Chem. 22 (2003) 657–669.